Faculty of Agriculture and Forestry
University of Helsinki

SENSITIVITY OF CEREAL YIELDS TO CLIMATE CHANGE: ESTIMATING RISKS AND EXPLORING ADAPTATION

Nina Pirttioja

DOCTORAL DISSERTATION

To be presented for public discussion with the permission of the Faculty of Agriculture and Forestry of the University of Helsinki, in Rasio Hall (ls B2), Forest Sciences Building, Latokartanonkaari 7, Helsinki, on 3 September 2021 at 13:00.

Helsinki 2021

Supervisors: **Professor Timothy R. Carter**
Finnish Environment Institute, Helsinki, Finland

Professor Annikki Mäkelä-Carter
Department of Forest Sciences,
University of Helsinki, Finland

Thesis committee: **Docent Frank Berninger**
Department of Forest Sciences,
University of Finland

Pre-examiners: **Professor Jørgen Eivind Olesen**
Department of Agroecology,
Aarhus University, Denmark

Senior Scientist Fredrik Wetterhall
Forecast Department, Evaluation Section,
Applications, European Centre for Medium-Range
Weather Forecasts (ECMWF), UK

Opponent: **Professor Paula Harrison**
UK Centre of Ecology and Hydrology, UK

Custos: **Professor Annikki Mäkelä-Carter**
Department of Forest Sciences,
University of Helsinki, Finland

Dissertationes Schola Doctoralis Scientiae Circumiectalis, Alimentariae, Biologicae

Publication No. 8/2021

ISBN 978-951-51-7298-3 (Print)
ISBN 978-951-51-7299-0 (Online)
ISSN 2342-5423 (Print)
ISSN 2342-5431 (Online)
Electronic publication at http://ethesis.helsinki.fi

Cover illustration: Nina Pirttioja

Unigrafia
Helsinki 2021

ABSTRACT

The agriculture sector is facing considerable challenges in ensuring food security under projected changes in climate and pressures to reduce its environmental impacts, among others. With changes in growing season and local growing conditions already being observed, adaptation is a key factor in aiming towards climate-smart, sustainable agriculture. Process-based crop models offer a tool for understanding complex interactions associated with crop, environment and management actions, and quantifying their impacts on various outputs. In the face of uncertainties associated with impact estimates, risk assessment has become an essential part of adaptation planning.

This study explored the use of a "scenario-neutral" approach for informing risk assessments in the context of crop production. Its main motivation was to examine novel insights offered by the approach for characterising uncertainties associated with modelled impacts compared to conventional scenario-based approaches, where impact estimates are tied to a given scenario. The approach utilises impact response surfaces (IRSs) to depict simulated period-mean sensitivities of cereal yields to systematic changes in baseline (1981–2010) temperature and precipitation. The analysis focused on sites in Finland, Germany and Spain, across a transect of contrasting environmental zones that hence facilitated an examination of the effect of site-specific growing conditions on the impacts of projected changes on cereal yields. The research encompassed a multi-model IRS study involving 26 crop models for wheat as well as an IRS study employing a single model for barley. In addition to analysing median responses of the model ensemble across the transect, approaches were developed for classifying and interpreting individual model responses. By combining IRSs with projections of climate interpreted probabilistically, likelihood of crop yield shortfall was estimated and its evolution throughout the 21st century visualised. This was estimated with a single crop model WOFOST for spring barley in Finland. Effects of adaptation on yield were considered through adapted sowing and cultivar choice. Evolution of future atmospheric carbon dioxide concentration [CO_2] defined by representative concentration pathways also used for climate projections (RCP4.5 and RCP8.5) was also considered when estimating likelihoods. With the multi-model ensemble study of wheat yield sensitivities [CO_2] was fixed at 360 ppm.

Simulated cereal yields were found to decline with warming and drying and increase with higher precipitation. The yield response in Finland was dominated by temperature. Precipitation change dominated the response of spring wheat in Spain, while the response was more mixed in Germany. The multi-model ensemble median response offered a consensus view of impact sensitivities, with individual model behaviour occasionally departing markedly from the average. IRS patterns across the multi-model ensemble

showed greater similarity in the pattern of modelled yield responses for Germany in comparison to Finland and Spain. Similarity in patterns was also associated with models of related genealogy. With respect to the effectiveness of tested adaptation options, results suggest that combining cultivars with short pre- and long post-anthesis phases with earlier sowing, offers most promise for obtaining the largest yield gains and smallest likelihoods of yield shortfall under future scenarios. Higher levels of [CO_2] generally compensate for yield losses with warming, with the effect emphasised with the biggest increases in temperature.

IRSs offer a valid alternative to conventional scenario-based approaches with many advantages for presenting and analysing results. IRSs can assist in model testing, comparison of results across models, studies and sectors and examination of various statistical characteristics of the response, greatly facilitated by the possibility to visually depict impact sensitivities in consistent ways. Use of multi-model ensembles with respect to both climate projections and crop impacts increases the robustness of results and provides information on the uncertainties around the yield estimates. The approach for estimating and visualising impact likelihoods provides improved understanding and transparency of concepts behind the likelihood estimates.

Keywords*:* adaptation, barley, climate change, crop model, ensemble, impact response surface, probabilistic projection, risk assessment, sensitivity analysis, wheat

ACKNOWLEDGEMENTS

This study was carried out at the Finnish Environment Institute (SYKE) as a collective effort of people and instances whom I acknowledge here. Financial support for carrying out the work, received from the Academy of Finland and the European Commission, is gratefully acknowledged. I want to thank CROPM, the crop modelling component of MACSUR (Modelling European Agriculture with Climate Change for Food Security), a project launched by the Joint Research Programming Initiative (JPI) on Agriculture, Food Security and Climate Change (FACCE) that provided the network essential for conducting the work done for the first two papers contributing to this thesis. With this, started my path towards finalising a PhD that is here finally presented.

First, I wish to express my deepest thanks to my supervisor Professor Tim Carter for all the things too many to list here that were crucial throughout the process of getting this done. None of this would have happened without your expertise and unwavering support. You have had to exercise a hefty amount of patience. Yet, with your scarce time, you were always willing to help. Though, at times, I bet, you tought that taking this on I regret. And in return, you only got some weak seedlings of leek. But now, momentarily, this is over and done and it is time to do something fun.

I thank Professor Annikki Mäkelä and Professor Frank Berninger for the guidance and support with the University side of things, as part of my advisory committee.

Many thanks go to my pre-examiners Professor Jørgen E. Olesen and Dr Frederik Wetterhall for the careful examination of my work and the valuable comments and suggestions given to improve the thesis. Thank you Professor Paula Harrison for agreeing to act as the opponent in my doctoral defence. I want to thank Annikki for acting as the custos at the event and Professor Juha Helenius for being part of the grading committee.

I am grateful to the large group of co-authors for their contributions and dedication to the work. Professor Reimund Rötter, your help with setting up the large network for the model intercomparison studies was invaluable and for that I am most grateful. Thank you, Dr Stefan Fronzek. You have been a wonderful colleague throughout the years, always willing to help and ever-patient throughout the turmoils every now and then experienced in ones life and developing a PhD. Dr Taru Palosuo, I want to thank you for always being so kind, positive and supportive and helping me with my numerous questions about the mysteries of crop models.

Professor Mikael Hildén, I sincerely thank you for providing the most wonderful and supportive atmosphere for working in the Climate Change Programme (ILMO) of SYKE. Similarly, I want to thank all my ILMO collagues past and present for the support and fun times that being part of this great group has offered. Anna L., an additional thank you for helping me with a particular translation task, my Finnish abstract. Further, with warm memories I recall the old third floor coffee table group and the discussions on nearly everything in life. I could surely now use the advise on renovating an old house. Numerous friendships were also formed that have lasted to this day. A special thanks goes to Anna R., Eerika, Ville-Veikko and Mikko.

Warmest thanks to my English School and associated friends, Elina, Emma, Jetta, Katja, Laura E., Laura R., Laura S. and Sonja. You have always been there, everything important in life I share with you and I know that I can always count on you. Thank you to Tuija and Oona for your friendship when I moved to a new town, Oulu. Your friendship throughout the geography studies that eventually led to me finding my way to my dream job in SYKE was invaluable.

Thank you mom and dad, Pirjo and Pentti. I didn't quite make it in time with this for you to see this happen. You always believed in me and were so proud of everything I did. I hope to succeed in extending the same level of love and support to my own sons. I want to also thank my parents-in-law, Pirjo and Rauno, for the relaxing times at the cottage, for always filling our fridge with berries and for helping out with our boys who were both born during this process.

Finally, my most sincere and heartfelt thanks to my boys, Teppo, Nooa and Topias. You mean the world to me. Teppo, your ability to solve and fix every weirdest thing never seizes to amaze me. Without your Python hotline during my internship WOFOST would have never started to run in batch and this thesis would have never happened. Though, just 2 years before I started this, when you defended yours, we swore that I would never in a million years start working on a PhD. Well, that really went according to plan. You have helped me through so much and with you I have had the greatest times and adventures of my life. Nooa and Topias, our two sweet, amazing sons. I still almost cannot believe how lucky we are to have such wonderful boys. I hope you never lose that happiness and curiosity towards all things big and small in life. I love you all so much.

CONTENTS

Abstract **3**

Acknowledgements **5**

Contents **7**

List of original publications **9**

Abbreviations **11**

1 Introduction **12**
- 1.1 Context and objectives ... 12
- 1.2 Climate change risk assessment ... 15
- 1.3 Climate change and crop production ... 17
  - 1.3.1 Observed changes in climate ... 17
  - 1.3.2 Climate change projections ... 18
  - 1.3.3 Impacts of climate change on crops ... 19
  - 1.3.4 Adaptation options for crop production ... 22
- 1.4 Modelling climate change impacts on crops ... 22
  - 1.4.1 Use of process-based models ... 22
  - 1.4.2 Model set-up and evaluation ... 25
  - 1.4.3 Uncertainties in crop model estimates ... 26

2 Material and methods **29**
- 2.1 Crop models ... 29
  - 2.1.1 Model ensemble ... 29
  - 2.1.2 The WOFOST model ... 31
- 2.2 Study sites ... 32
- 2.3 Crop and soil data ... 33
- 2.4 Climate data ... 35
  - 2.4.1 Observed climate ... 35
  - 2.4.2 Future projections ... 35
- 2.5 Impact response surfaces and their applications ... 37
  - 2.5.1 Construction of impact response surfaces ... 37
  - 2.5.2 Measures for analysing model ensemble results ... 38
  - 2.5.3 Classification of crop model responses ... 38
  - 2.5.4 Estimation of impact likelihoods ... 39

3 **Results** **41**
3.1 Model performance under baseline conditions 41
3.2 Sensitivity of cereal yields to temperature and precipitation in Europe 44
3.3 Identification and interpretation of differences in simulated yield responses 49
3.4 Approaches to estimating likelihood of yield shortfall throughout the 21st century 52
3.5 Cereal yield responses to adaptation options under a changing climate 55

4 **Discussion** **58**
4.1 Model calibration and evaluation 58
4.2 Uncertainties in crop yield estimates 60
4.3 Sensitivity of cereal yields to climate change 63
4.4 Likelihoods of yield impacts and effect of adaptation options 65
4.5 Modelling set-up 67
4.5.1 Choice of model 68
4.5.2 Utility of the IRS-based approach 69
4.5.3 Implications of refinements in the modelling set-up 73
4.6 Outlook 74

5 **Conclusions** **76**

References **79**

Appendix 1. Criteria for classifying model complexity (TABLE 2) **93**

Erratum **94**

Original publications **95**

LIST OF ORIGINAL PUBLICATIONS

This thesis synthesises the research reported in the following three original publications, which are referred to in the text by their Roman numerals:

I **Pirttioja, N**., Carter, T.R., Fronzek, S., Bindi, M., Hoffmann, H., Palosuo, T., Ruiz-Ramos, M., Tao, F., Trnka, M., Acutis, M., Asseng, S., Baranowski, P., Basso, B., Bodin, P., Buis, S., Cammarano, D., Deligios, P., Destain, M.-F., Dumont, B., Eweret, F., Ferrise, R., François, L., Gaiser, T., Hlavinka, P., Jacquemin, I., Kersebaum, K.C., Kollas, C., Krzyszczak, J., Lorite, I.J., Minet, J., Minguez, M.I., Montesino, M., Moriondo, M., Müller, C., Nendel, C., Öztürk, I., Perego, A., Rodríguez, A., Ruane, A.C., Ruget, F., Sanna, M., Semenov, M.A., Slawinski, C., Stratonovitch, P., Supit, I., Waha, K., Wang, E., Wu, L., Zhao, Z., Rötter, R.P. 2015. Temperature and precipitation effects on wheat yield across a European transect: a crop model ensemble analysis using impact response surfaces. *Climate Research* 65:87-105.

II Fronzek, S., **Pirttioja, N**., Carter, T.R., Bindi, M., Hoffmann, H., Palosuo, T., Ruiz-Ramos, M., Tao, F., Trnka, M., Acutis, M., Asseng, S., Baranowski, P., Basso, B., Bodin, P., Buis, S., Cammarano, D., Deligios, P., Destain, M.-F., Dumont, B., Ewert, F., Ferrise, R., François, L., Gaiser, T., Hlavinka, P., Jacquemin, I., Kersebaum, K.C., Kollas, C., Krzyszczak, J., Lorite, I.J., Minet, J., Minguez, M.I., Montesino, M., Moriondo, M., Müller, C., Nendel, C., Öztürk, I., Perego, A., Rodríguez, A., Ruane, A.C., Ruget, F., Sanna, M., Semenov, M.A., Slawinski, C., Stratonovitch, P., Supit, I., Waha, K., Wang, E., Wu, L., Zhao, Z., Rötter, R.P. 2018. Classifying multi-model wheat yield impact response surfaces showing sensitivity to temperature and precipitation change. *Agricultural Systems* 159:209-224.

III **Pirttioja, N**., Palosuo, T., Fronzek, S., Räisänen J., Rötter, R. and Carter, T.R. 2019. Using impact response surfaces to analyse the likelihood of impacts on crop yield under probabilistic climate change. *Agricultural and Forest Meteorology* 264: 213-224.

Author's contribution to the publications

I Pirttioja was the lead author. Rötter and Carter proposed the original idea of this paper. Pirttioja designed this large multi-model study together with Carter, Fronzek and Rötter, was in charge of designing protocols for and overseeing the model runs performed by individual modelling groups, compiling and analysing the results and writing the paper, assisted by a core team of authors and an extended team representing all modelling groups.

II Pirttioja was partly responsible for designing the study, which employed the same data collected and compiled by Pirttioja for paper I. Fronzek, Carter, Rötter and Pirttioja developed the concept for the paper. Pirttioja was responsible for designing, setting up and performing the calculations using the expert diagnostic approach for classifying impact response surfaces. She also participated in analysing the results and writing the paper. Fronzek was lead author of the paper.

III Pirttioja was the lead author. She designed the study with Carter, who also provided the original idea. Palosuo and Rötter provided insights on the agricultural dimensions of the study and Fronzek on the methods relating to estimating impact likelihoods. Räisänen provided the data on probabilistic climate projections. Pirttioja performed all the model runs, compiled and analysed the results and was responsible for writing the paper.

ABBREVIATIONS

ANN	Annual
CMIP	Coupled Model Intercomparison Project
[CO_2]	atmospheric carbon dioxide concentration
CV	coefficient of variation
DE	Germany
DOY	day of the year
EDA	expert diagnostic approach
ES	Spain
FACE	Free Air CO_2 Enrichment
FAO	Food and Agriculture Organization of the United Nations
FI	Finland
GCM	global climate model
GIS	Geographic Information Systems
IQR	inter-quartile range
IRS	impact response surface
LAI	leaf area index
ppm	parts per million
RCP	Representative Concentration Pathway
SDA	statistical diagnostic approach
SRES	Special Report on Emissions Scenarios
TSUM	thermal time needed to reach a certain phenological stage

1 INTRODUCTION

1.1 CONTEXT AND OBJECTIVES

Agriculture is the major use of land across the globe and a major economic, social and cultural activity, providing a wide range of ecosystem services. Croplands supply a large part of the food caloric and protein supply for global consumption (FAOSTAT, 2018), resulting in crop production being of vital importance for society. However, crops are highly sensitive to variations in weather and climate. The overwhelming evidence that human activities have been changing the climate and are projected to cause even larger changes in the future (IPCC, 2014c) provides strong motivation for assessing the potential impacts, associated risks and possible benefits of projected climate changes for crop production (Howden et al., 2007).

While having to meet the demands of feeding a growing global population, including the aim of ending hunger by 2030, goal 2 of the United Nations Sustainable Development Goals, considerable transitions in the production of food and the agricultural sector in general are needed to meet the Paris target of limiting global warming to 1.5°C (IPCC, 2018). Stringent mitigation targets suggest, for example, a requirement to decrease cropland for food and feed production under most 1.5°C pathways (IPCC, 2018). At the same time reductions in the use of fertilisers are being enforced through different agricultural policies to reduce the environmental impacts of agriculture. Consequently, considerable intensification of agricultural productivity is needed to ensure food security and achievement of associated goals, but with less external inputs and under more constrained conditions than earlier for resources such as water availability and agricultural land (Smith et al., 2019).

Alongside mitigation targets and actions, adaptation is a key factor in shaping the future of climate change impacts on food production. Options ranging from incremental coping measures implemented at the local level (e.g. changes in sowing dates) to transformational solutions affecting the system on a more profound level through marked changes towards something new (e.g. change in business scale, structure, and location – Panda, 2018) all play a strong role in improving the climate resilience of the agricultural sector and aiming towards sustainable risk management and climate-smart agriculture (Lipper et al., 2014). Through the multitude of uncertainties associated with climate change impact estimates, risk assessment has become a core part of adaptation decision making (e.g. Hinkel et al., 2013; Willows and Connell, 2003) and a recommended procedure for improving climate change impact assessments (White et al., 2011).

The European Green Deal, announced in December 2019, provides a roadmap for making the European Union's economy sustainable by aiming to turn the challenges posed by climate and environmental changes into opportunities (European Commision, 2019). As part of this, the Commission adopted a new Climate Adaptation Strategy in early 2021, building on the 2018 evaluation of the earlier 2013 adaptation strategy. Its principle objectives are: to make adaptation smarter, swifter and more systematic, and to step up international action on adaptation to climate change by promoting sub-national, national and regional approaches to adaptation (European Commision, 2021). In Finland, the first national adaptation plan extending to 2022 was approved in 2014 (MMM, 2014), updating the National Adaptation Strategy 2005. The Climate Change Act (YM, 2015) that soon after entered into force, states that a national adaptation plan should be approved at least every ten years. While implementing the EU Strategy on Adaptation to Climate Change in Finland, the plan aims to ensure that Finland has the capacity to manage the risks associated with climate change and adapt to those changes. The objectives for achieving this are "1) adaptation has been integrated in the planning and activities of both the various sectors and their actors, 2) the actors have access to necessary climate change assessment and management methods and 3) research and development work, communication and education and training have enhanced the adaptive capacity of society, developed innovative solutions and improved citizens' awareness on climate change adaptation".

Process-based models can be valuable tools for helping to understand the complexities of climate change impacts on global food production as well as issues associated with mitigation and adaptation. A lot of research on the topic has been carried out around the world using diverse approaches, scenarios and models over the recent decades. While originally the emphasis was on model development (e.g. de Wit, 1958; Duncan et al., 1967) and description of potential adverse effects associated with climate change impacts (e.g. Rosenzweig and Parry, 1994), only rather recently the focus has shifted towards informing adaptation and finding solutions, with the amount of research on the topic now vast and growing (e.g. Challinor et al., 2014; Olesen et al., 2011; Ruiz-Ramos et al., 2018; Rötter et al., 2018a). As the majority of models were developed already much earlier, the need has been highlighted to improve these by incorporating the latest knowledge about crop responses to changes in climate, through the use of ensemble methods for quantifying uncertainties and by generating good quality field data for model development, testing and application (Rötter et al., 2011a). As a response to these requirements, global model intercomparison programmes were launched, offering frameworks for the consistent projection of climate change impacts, with the aims of informing model improvement efforts and quantification of uncertainties associated with impact estimations, among others (Rosenzweig et al., 2013; Schellnhuber et al., 2014).

The main motivation for this study is to contribute to these efforts by exploring and developing methods for assessing crop responses to climate change that may more readily offer insights about uncertainties in impact estimates than conventional approaches. The common "scenario-based" approach models impacts that are conditional on a given scenario of future climate and related conditions, and uncertainty around estimates is commonly quantified only to a limited extent via a selection of scenarios selected to embrace a range of uncertainties. The approach used in this study utilises "scenario-neutral" impact response surfaces (IRSs) constructed from crop model simulation results for cereals, visually depicting the sensitivity and behaviour of modelled responses (e.g. of grain yield) across a range of plausible changes in key driving variables. Scenarios are also applied, but at a later stage, to estimate likelihoods of yield impacts. While the importance of mitigation is acknowledged, the main focus of the study, however, is on climate change impacts and adaptation.

The IRS approach is applied to address specific objectives of:

1. examining the performance of crop models and sensitivities of cereal yields under present-day conditions and assumed changes in temperature and precipitation (Papers **I** and **III**),
2. identifying differences in model behaviour across a large ensemble of wheat models and in different parts of Europe and possible reasons for such differences (Papers **I** and **II)**,
3. estimating the likelihoods of future yield impacts on barley in Finland by combining IRSs with projections of climate change interpreted probabilistically and accounting for the direct effects of changing atmospheric carbon dioxide concentrations (Paper **III**),
4. analysing the effectiveness of potential farm-level adaptation options (Paper **III**),
5. exploring the implications of using different approaches for applying models and analysing their results (Papers **II** and **III**).

The papers contributing to the thesis are organised such that paper **I** introduces the method of constructing IRSs that is the basis for all subsequent analyses. It illustrates the use of IRSs for examining aspects of cereal yield sensitivities and model stability under perturbed climates. The focus is on the ensemble median responses of wheat (spring and winter varieties) from a large model ensemble at sites across Europe. Paper **II** uses outcomes from the ensemble of model simulations in paper **I** to look in more detail at individual model responses. It introduces two new approaches for classifying patterns of yield responses that aid in identifying differences between models and study sites and in seeking possible reasons for such differences. Paper **III** extends the application of IRSs by combining them with projections of climate change interpreted probabilistically. The method introduced in the paper facilitates the estimation of likelihoods of impacts on cereal yields and an assessment of the effectiveness of adaptation options. In contrast to the first two papers, the method is illustrated and tested for spring barley with a single crop model, WOFOST, at a site in Finland. Specific attention is paid to the implications for study results of applying alternative modelling approaches.

1.2 CLIMATE CHANGE RISK ASSESSMENT

The concept of risk has increasingly become central in assessments of climate change impacts and vulnerability and in the design of potential adaptation and mitigation measures in the decision-making context (e.g. Hinkel et al., 2013; IPCC, 2014c; Tuomenvirta et al., 2018; Willows and Connell, 2003). Typical decision frameworks comprise several stages from identifying adaptation needs and options, achieved through risk assessment, through appraisal of the adaptation options to implementation and monitoring of adaptation actions. It is an iterative learning process designed to cover the whole decision-making cycle, allowing decision-makers to define and refine their decisions in the face of risk (e.g. Hinkel et al., 2013; Willows and Connell, 2003). The principle reason for linking adaptation and risk assessments is to provide an understanding of how to frame decisions that bridge between the known present and the unknown futures associated with a multitude of uncertainties (Jones and Mearns, 2005).

Fundamentally, risk is defined as the product of the probability of some event or sequence of events occurring and the consequences of that event (Risk = Probability x Consequence – e.g. Lavell et al., 2012). In the context of climate change, an approach often used is to express these concepts by defining risk as a function of hazards, exposure and vulnerability where all three can be thought of as potentially affecting either the probability of an event or its consequence depending on the context (IPCC, 2014c).

In this formulation, hazard refers to climate-related events and trends (i.e. changes in the mean and/or variability) that may induce adverse impacts such as loss or damage to livelihoods and environmental resources. This can also be interpreted to refer to related environmental variables such as atmospheric CO_2 concentration [CO_2], tropospheric ozone or other air pollutants. Exposure is described as the presence of the entity of interest, such as people, livelihoods, species or resources, in places and settings that could be adversely affected by the hazard. Vulnerability is defined as the propensity or predisposition of the exposed entity to be adversely affected, encompassing concepts such as sensitivity, susceptibility and lack of capacity to cope and adapt (IPCC, 2014a). These three different components that constitute risk are driven by changes in the climate system and socioeconomic processes. Through changes in natural variability and consequences of the anthropogenic contribution that can be reduced through climate change mitigation, climate affects hazards that interact with vulnerabilities and exposure. On the other hand, various socioeconomic processes such as changes in economic activity, population or land use and actions aiming towards adaptation or mitigation may affect and alter exposure and vulnerability (IPCC, 2014c).

As to the approach chosen for conducting a risk assessment, a variety of alternatives can be applied depending on the focus and emphasis of the assessment. One possible way of characterising the available options is to distinguish between a hazard oriented and a vulnerability-exposure oriented approach (Jones and Mearns, 2005). The distinction between the two results from emphasising impacts either from the point of the main climate drivers undergoing change (the hazard) or rather from the perspective of the socioeconomic context in which the impacts are occurring, which are also changing. In the *hazard oriented approach,* a level of a specific hazard is first fixed, such as a threshold temperature of 47.5°C considered lethal for wheat (Porter and Gawith, 1999). Risk is then defined, for example, as the percentage of years when the crop dies due to exceedance of the lethal temperature limit during critical growth phases under scenarios of future climate. Vulnerability and exposure are not treated separately but rather taken as inherent features of the defined threshold with a consequence of crop mortality. In the *vulnerability-exposure oriented approach* a threshold is defined with respect to a specific impact or outcome, such as the minimum level of yield necessary for ensuring a viable livelihood. By applying an impact model under conditions defined by future scenarios, the risk of falling short of the threshold can be assessed. Here, vulnerability and exposure of the crop are pre-determinants of the impact, which are case-specific and may be due to a variety of possible hazards acting with other factors throughout the growing season of the crop.

The two approaches are complementary and can be used separately or in combination to inform of different aspects of risk (Jones and Mearns, 2005; Lavell et al., 2012). In this study, IRSs are adopted for assessing risks, largely following the logic of the vulnerability-exposure oriented approach. They are constructed from the results of a sensitivity analysis of a crop model and used to analyse crop yield responses to projected changes in climate. As such, the assessment of vulnerability (of yield) is separated from the hazard (posed by climate), making the appraisal of vulnerability independent of the evolution of climate models and downscaling schemes (see sub-section 2.5.1 for details; Keller et al., 2019).

It should be noted that in the papers included in this study, the term "likelihood" is used to refer to specific instances associated with the concept of risk, namely the likelihood of falling short of a defined yield threshold. The decision was made to use "likelihood" instead of risk in part to recognise that risk focuses on adverse impacts whereas likelihood is a neutral term that also encompasses positive impacts. Further, risk was regarded here as a broader concept where assessments of risk could theoretically be based on a variety of measures, approaches and estimates, instead of a single measure associated with yield levels. In this context, "likelihood" should also be understood to relate to both sides of the general risk equation, including both the probability of an event occurring and its importance through the associated consequences, rather than being synonymous with the concept of probability used to describe some of the climate projections (see sub-section 2.4.2).

1.3 CLIMATE CHANGE AND CROP PRODUCTION

Throughout the history of organised agriculture, crop production around the world has fluctuated from year to year in response to weather during the growing season (Monteith, 1981). This can be seen as variations in the productivity of the crop (production per unit area cultivated), commonly reported as crop yield (describing the harvestable component of the crop, such as the grain in cereals), though weather can also affect production by limiting the area cultivated. Due to the natural variability of climate, agriculture has throughout its existence had to adapt in reaction to changing conditions. However, due to human activities changing the emissions of important greenhouse gases (CO_2, methane [CH_4], and nitrous oxide [N_2O]) and aerosols as well as altering the reflectivity of the land through changes in land surface properties, the Earth's radiation budget is being perturbed, producing a radiative forcing that affects climate. The concept of radiative forcing describes the net change in Earth's energy balance in response to an external perturbation. There is growing evidence that due to human influence on radiative forcing, climate is now changing at unusual levels and rates, posing novel risks often outside the range of experience (Bindoff et al., 2013; Cubasch et al., 2013).

1.3.1 OBSERVED CHANGES IN CLIMATE

Of the many indicators of a changing climate, increase in mean surface temperature, changes in precipitation and increase in [CO_2] are among the most relevant for agriculture and crop production. Global mean surface temperature has increased during the past century and since the 1980s each decade has been significantly warmer than any preceding decade in records dating back to the 1850s (Hartmann et al., 2013). In large parts of Europe warming has been accompanied by an increase in the frequency of heatwaves and hot days (Kovats et al., 2014). For precipitation, the signals are more mixed. However, observations suggest that in northern Europe annual precipitation has increased since 1950, while in parts of southern Europe it has decreased (Kovats et al., 2014). For Europe, there is also evidence of an increase in the observed frequency or intensity of heavy precipitation events (Hartmann et al., 2013). Finally, the abundance of [CO_2], as an annual average, has increased by 49% from 277 ppm in 1755 (IPCC, 2013) to 414 ppm in 2020[1].

[1] Source: Global Greenhouse Gas Reference Network. Dr. Pieter Tans, NOAA/ESRL (www.esrl.noaa.gov/gmd/ccgg/trends/) and Dr. Ralph Keeling, Scripps Institution of Oceanography (scrippsco2.ucsd.edu/). Data accessed: 25 March 2021. https://www.esrl.noaa.gov/gmd/webdata/ccgg/trends/co2/co2_annmean_mlo.txt

1.3.2 CLIMATE CHANGE PROJECTIONS

Projections of future climate change are typically obtained from climate model simulations driven by scenarios of external forcing. Global climate models (GCMs), representing physical components of the climate system (atmosphere, ocean, land and sea ice) as well as various biogeochemical cycles (e.g. of carbon and sulphur) are the most advanced and comprehensive tools available for simulating the response of the global climate system to external forcing and making projections of future climate. It should be noted that GCM is used here to refer to a range of models of varying complexity from Atmosphere-Ocean General Circulation Models to more complex Earth System Models. All are represented in the Coupled Model Intercomparison Project Phase 5 (CMIP5 – Taylor et al., 2012). Most climate projections applied in this study were derived from CMIP5 model simulations.

Representative Concentration Pathways (RCPs) are time-dependent trajectories of greenhouse gas and aerosol concentrations used in CMIP5 that cover a wide range of possible magnitudes of radiative forcing by 2100. Originally, four RCPs were specified : RCP2.6 (lowest), RCP4.5 (medium-low), RCP6.0 (medium-high) and RCP8.5 (highest) (van Vuuren et al., 2011). Depending on the RCP, global mean temperature change across the CMIP5 models (5-95% projection range) varies from 0.3-1.7 °C (RCP2.6) to 2.6-4.8 °C (RCP8.5) by 2081-2100 relative to 1986-2005. However, there is regional variation in the projected changes with the Arctic region being projected to warm most. Occurrences of various high temperature events are projected to increase in frequency, magnitude and duration. Along with the increase in global mean temperature, global precipitation is also projected to increase. For scenarios other than RCP2.6 the projected increase is 1 to 3% $°C^{-1}$ by the end of the century, again with substantial spatial variation. In Europe, the mid to high latitudes are projected to get wetter, especially in winter, while drying is projected for lower latitudes in the south, including the Mediterranean region. Generally, the contrast between dry and wet regions is estimated to increase. With the rising temperature, a shift to more intense individual precipitation events is projected globally. Increasing annual surface evaporation is likely to decrease soil moisture, resulting in a higher risk of agricultural drought in presently dry regions such as the Mediterranean (Collins et al., 2013). Projections of [CO_2] at the end of the 21st century range from 421 ppm (RCP2.6) to 936 ppm (RCP8.5. – IPCC, 2013), the latter implying a further increase of 126% from the level of 2020 (414 ppm)[1]. Note, that since CMIP5, an even lower forcing has been introduced, represented by RCP1.9, to inform of actions required to meet the 1.5°C Paris Agreement goal (O'Neill et al., 2016).

1.3.3 IMPACTS OF CLIMATE CHANGE ON CROPS

Since the introduction of improved crop management practices and cultivars with higher yielding potential in the 1960s, crop yields increased steadily nearly worldwide for decades (Evans, 1996). Yet, since the early to mid-1990s, a stagnation of yields has been experienced in Europe. Although this may largely be explained by changes in agriculture and environmental policies, there is also evidence of climate trends accounting for part of the stagnation (e.g. Brisson et al., 2010; Moore and Lobell, 2015; Peltonen-Sainio et al., 2009b).

For understanding the effects of climate on crops, it can be useful to distinguish between two features of crop response: development and growth. Development is here defined as the succession of phenological stages, describing the timing of crop life cycle events, from the initiation of growth to the death or harvest of the plant. Growth, on the other hand, refers to the quantitative increase in size, mass or volume of a crop or its parts. While growth is mostly associated with capture and allocation of resources (light, water and nutrients), temperature is the most important factor affecting the rate of plant development (Sadras et al., 2016). Crop species differ in their responses to temperature throughout their life cycle, with each species having a range of minimum and maximum temperatures that limit development and observable growth (Hatfield and Prueger, 2015). For example, for wheat the optimum range is from 17 to 23°C over the course of the entire growing season, with the lowest limit at 0°C and highest at 37°C, beyond which development stops (Porter and Gawith, 1999). Cultivars grown at each location have normally been bred to develop and mature under ambient conditions such that they produce the best possible yield. With warming, development is accelerated, which shortens the duration of phenological stages, hence allowing less time for the plant to capture light, water and nutrients. Consequently, biomass production is decreased, the grain-filling period is reduced, and crops mature earlier, which is likely to result in reduced final yields. Additionally, repeated exposures to extreme temperatures at critical phenological stages, in particular at the reproductive stage, can have detrimental effects on the yield potential (Hatfield and Prueger, 2015; Lawlor and Mitchell, 2000).

Possible benefits of warming include lengthening of the growing season (period during which temperature and moisture conditions are suitable for crop growth) in the middle and higher latitudes of Europe where a short growing season is one of the main limiting factors to crop production. For cereals in Finland, it has been found that between 1965 and 2007 the start of the growing season (defined as the period when daily mean temperatures permanently exceed 5°C in the spring) had already advanced by 0.6 to 1.7 days per decade, depending on the region (Kaukoranta and Hakala, 2008). Lengthening of the season can enhance yield potential through a longer development time and offers the possibility of growing new, more productive crops and cultivars (Olesen and Bindi, 2002). Already, during 1996–2016, farmers have been reported to be adopting later maturing cultivars more frequently than earlier (Peltonen-Sainio and Jauhiainen, 2020). However, warming autumns have less of an effect for additional cereal yield gains as even the longest developing cultivars, if sown earlier, would be likely to mature around the same time or earlier than current cultivars. Further, various risks to late season harvests associated with projected elevated autumn precipitation may limit the benefits of warming autumns. One possible option may be provided through double-cropping of a primary and a cover crop, with several possible benefits provided through e.g. biomass production, soil cover and increase in soil carbon (Peltonen-Sainio et al., 2018).

While temperature has a profound effect on plant development, precipitation is a major climatic determinant of crop productivity, as the primary source of soil moisture. In non-irrigated conditions, much of crop production is water-limited to some extent. With warming evaporative demand in increased, which may result to water stress and consequently to declining potential for net primary production (Turral et al., 2011). To prevent dehydration crops control water loss by stomatal closure, which also limits the uptake of CO_2 and thus growth (Morison et al., 2008). Greater precipitation generally increases yields by reducing water stress caused by water shortage. However, excessive water can lead to declines in yield through waterlogging, which causes oxygen shortage in the soil. Excess moisture can also promote pest and pathogen infestation, while high levels of soil moisture may hinder field operations, potentially delaying or even preventing sowing and/or harvest. Episodes of intense rainfall may damage young plants, promote lodging, cause soil erosion and increase the potential for flooding (Kristensen et al., 2011; Rosenzweig et al., 2001; Sutherst et al., 2011).

The increasing level of [CO_2] is known to increase plant biomass and yield, particularly in crops relying on the C_3 pathway for carbon fixation, such as barley and wheat, by increasing light-use efficiency and rates of photosynthesis and through enhanced water use efficiency resulting from reduced stomatal conductance and transpiration (Drake et al., 1997). However, the beneficial effects of elevated [CO_2] have been found to be reduced or even negative with insufficient nutrients (Ainsworth and Long, 2005). Thus, the presumption is that adequate nitrogen is available to sustain the increased assimilation and for maintaining an optimal carbon and nitrogen ratio, critical for various growth and development processes governing yield and seed quality (Kant et al., 2012). These direct CO_2 effects on crops can be expected to interact with changes in climate. For example, although the benefits of elevated [CO_2] are usually relatively larger in warmer environments, the combination of a doubled [CO_2] and warming of 1.6 to 4.0°C has been found typically to reduce yields (Amthor, 2001). However, due to the complexities of interactions between temperature, nutrients, surface ozone (O_3), pests and weeds, which are not well understood, uncertainty remains in the magnitude of the CO_2 effect (Soussana et al., 2010).

To summarise, projected changes in climate can be expected to alter regional agro-climatic conditions in Europe. Table 1 attempts to capture some of the key climate changes and their likely impacts. For example, in northern Europe, in addition to a potential lengthening of the growing season, the conditions for rainfed production may improve. Conversely, in some of the currently high yielding regions of Europe as well as the Mediterranean region, the suitability for rainfed crop production may decline, limiting crop growth unless irrigation is applied (Trnka et al., 2011, and see Table 1).

Table 1. *Observed and projected climate changes of relevance for agriculture and some key impacts on crops across the main regions of Europe (revised extract from Füssel et al., 2017).*

Indicator/impact domain	Variable	Northern		Temperate		Southern	
		Boreal and Arctic		Continental		Mediterranean	
		Obs	Proj	Obs	Proj	Obs	Proj
Changes in climate							
Global and European temperature	Temperature	↗		↗		↗	
Heat extremes	Frequency of warm days/heat waves	↗		↗		↗	
Mean precipitation	Annual precipitation	↗		↗↘	↗→	↘→	
Heavy precipitation	Intensity	↗		○	↗	↘→	↗↘
Climate impacts on crops							
Growing season for crops	Duration	↗		↗		↗→	
Agro-phenology	Day of spring events[1]	↘		↘		↘	
Water-limited crop yield	Average yield	○	↗	○	→	○	↘
	Adverse climatic conditions	↗		↗→		↗→	
Crop water demand	Water deficit	○		↗↘		↗↘	↗→

Legend (changes in blue are beneficial; red are adverse; black are neither beneficial nor adverse or changes are small)

↗	Increase throughout most of a region	Dominating trend in at least two-thirds of a region; opposing trend in less than 10%
↘	Decrease throughout most of a region	
↗→	Increase in substantial parts of a region	Trend in between one-thirds and two-thirds of a region; opposing trend in less than 10%
↘→	Decrease in substantial parts of a region	
↗↘	Increases as well as decreases in a region	Trends in both directions in at least 10% of a region
→	Only small changes	Trends in either direction in less than 10% of a region
○	No data available in Füssel et al. (2017)	

Notes: [1] Day = day number from 1 January; decrease indicates that the day is earlier; Obs = observation/past trend; Proj = model projection; agreement between Obs and Proj is indicated by one centred arrow; judgements are based on information for varied time periods, emissions scenarios and socio-economic scenarios.

1.3.4 ADAPTATION OPTIONS FOR CROP PRODUCTION

Due to the scale and magnitude of potential climate change impacts on crop production and food security, identifying and evaluating options for adaptation has become a critical concern. The purpose of adaptation is to avoid or reduce the negative effects of potential climate risks over the coming decades as well as to benefit from possible positive effects, with economic considerations providing strong motivation for research and implementation (Howden et al., 2007). The main types of adaptation can be classified into incremental and transformational adaptation. In incremental adaptation the aim is to maintain the essence and integrity of the existing system or a process. Options include altering the timing of cropping activities such as sowing and harvest, introduction of new crop varieties, adjustments in the use of fertilisers and pesticides and more efficient water management. In contrast, transformational adaptation involves profound changes in the system and its attributes that are often adopted at a larger scale than incremental adaptation measures. Examples include diversification of crops and activities even outside agriculture leading to changes in land use, investments in new technologies and development of infrastructure, information and engagement processes (Noble et al., 2014). Results across a large set of studies suggest that adoption of adaptation measures, either individually or in combination, could substantially offset negative impacts of climate change as well as enhance the benefits of positive impacts (Porter et al., 2014b).

1.4 MODELLING CLIMATE CHANGE IMPACTS ON CROPS

The techniques available for simulating crop responses to climate can be roughly categorised into statistical and process-based models. Statistical models, used for example for crop yield prediction, are based on relatively simple regression equations calibrated with historical yields and simple measurements of weather (Lobell and Burke, 2010). However, when estimating the impacts of a changing climate on crop yields, the reliance on statistical coefficients alone, instead of describing the fundamental biophysical mechanisms governing crop responses to climate factors, may limit their usability. Under future climate change past relationships may not hold and a capability is needed to extrapolate beyond the range of historically observed climate and crop information (Rosenzweig and Hillel, 1998).

1.4.1 USE OF PROCESS-BASED MODELS

Process-based based models offer an alternative by formulating mathematical descriptions of known or hypothesised physical, chemical, and biological processes, such as those of photosynthesis (e.g. Farquhar et al., 1980) and water movement in soils (e.g. Richards, 1931), reflecting an understanding of the phenomena being simulated. Dynamic, process-based crop models, widely used in climate change research and also applied in this study, simulate and quantify the development and growth of a crop throughout the growing season in interaction with its environment (Asseng et al., 2014b). The approach applied simulates changes in quantifiable state variables such as biomass, grain yield or the amount of water in the soil in response to external drivers such as climate variables. The states are further associated with rate variables characterising their rate of change at a certain instant in response to specific processes (de Wit, 1982). Such models are capable of quantifying the interactions between crop genotype, environment and management actions and their impacts on various outputs (Rötter et al., 2018b). The origins of process-based crop models can be traced back to the 1960s to the early works of de Wit (1965) and Duncan et al. (1967) on modelling canopy photosynthesis. Current applications deal with a variety of societally relevant issues, such as crop yield forecasting (e.g. Rembold et al., 2019), climate change impacts and adaptation (Ewert et al., 2015) and food security (Wheeler and von Braun, 2013). They are also necessary components of more complex agricultural system models, which combine crop, livestock and farming system models and are needed to support efforts to ensure that the food demand of the next 50 to 100 years is met in a sustainable way, both environmentally and economically (Jones et al., 2017).

A multitude of process-based crop models, referred to hereinafter simply as crop models, has been developed since the 1960s, distinguished by differences in the level of detail and methods that are used in the models to represent key bio-physical processes and components (e.g. Rivington and Koo, 2010). Some key processes are illustrated in Figure 1. Common processes include the phenological development of the crop from sowing to ripening, included in all models as a function of accumulated temperature, and often adjusted to account for day-length and (for over-wintering crops) a winter-chilling requirement (vernalisation). Light interception and utilisation along with allocation of accumulated biomass to different crop organs further describe the growth and development of the crop and its components. Water availability, a key resource for crop growth, is defined through representations of soil water dynamics and evapotranspiration. Processes describing nutrient dynamics are also often included in the models. Various other processes associated with effects of CO_2 and different stresses on the crop may also be described. As a result, the models are often very complex, containing multiple variables and parameters. Modelling of the effects of pest and diseases (Asseng et al., 2014b; Donatelli et al., 2017) and the impacts of excessive rainfall and other damages (e.g. lodging, frost and flooding) still largely remain a challenge for the scientific community and are often not included in crop models such as the ones applied in this study (Rivington and Koo, 2010).

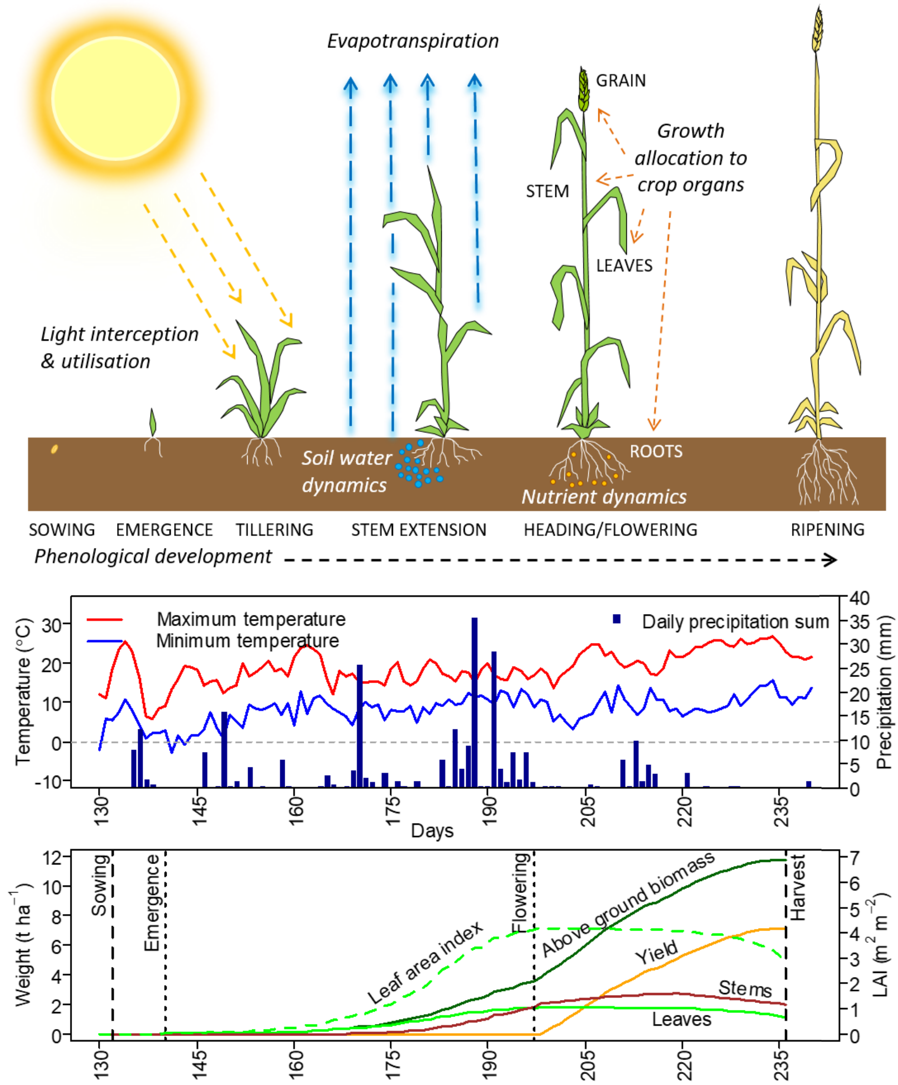

Figure 1 Top panel: Phenological stages of a typical spring-sown cereal with illustrations of some key processes required for modelling crop growth and yield (in italics). Middle panel: Minimum and maximum daily temperature and daily precipitation sum during the growing season of an example year (1996) in Jokioinen, Finland. Bottom panel: Crop model (WOFOST) output based on the same weather data, illustrating the key stages of growth from sowing to harvest (vertical dashed lines) and the development of leaf area index (LAI) and different crop components throughout the growing season.

Depending on model structure and the level of detail in the process descriptions, varying amounts of input data on weather, soil and management are required for running the models. Models are usually run on a daily time step, requiring for example temperature and precipitation data at the same resolution (Figure 1, middle panel). Model output typically describes the daily states of crop development and amounts of total above ground biomass and its partitioning to different crop components (e.g. grain yield, stems and leaves) during the growing season, among other issues related to crop growth and use of resources. Figure 1, bottom panel, illustrates an example of a run of daily output from the crop model WOFOST for an individual growing season in south-western Finland, corresponding with the weather depicted above.

1.4.2 MODEL SET-UP AND EVALUATION

Model calibration or parameter value estimation involves adjusting the values of influential model parameters in order to improve the fit between simulated results and measured data. The procedure may need to be repeated many times with the same model, especially when it is being applied across diverse conditions. As such, it can have a large effect on model outcomes, as different parameter values operating on the same model equations can give very different results, both qualitatively and quantitatively. Typically, values of most model parameters are fixed based on studies of individual model processes. However, if parameter values for certain processes cannot be obtained independently (e.g. due to a lack of observations), a combination of expert judgement to select key parameters and use of appropriate statistical methods to estimate their values is recommended (Wallach et al., 2014). There are no clear guidelines or a standard approach applied as to which parameters should be chosen for calibration or how the calibration should be performed. For more information on alternative calibration methods see (e.g. He et al., 2010; Li et al., 2015; Ramirez-Villegas et al., 2017; Seidel et al., 2018; Tao et al., 2009; Wallach et al., 2014).

Model evaluation. An issue linking closely with model calibration is model evaluation, interpreted here as the comparison between observed and simulated values. The aims of model evaluation serve to provide understanding of the range of conditions the model is applicable in, inform of needs for model improvement and attach confidence in the model results produced. A further consideration is the difference between theoretical yield levels (often portrayed by simulation results) and actual farmers' yields, known as the yield gap. When performing crop model simulations with water and nutrients non-limiting and assuming no biotic stresses affecting yields, the resulting yields can be used as an estimate of yield potential, providing a yield ceiling for a given crop and location (van Ittersum and Rabbinge, 1997). For irrigated systems and humid climates with adequate water supply, yield potential is considered the most relevant benchmark. For rainfed crops, it is more appropriate to consider water-limited potential yields as the highest achievable level. In reality, actual farmers' yields rarely approach these potential levels, as this is unlikely to be either economically or practically feasible, nor even necessary or desirable due to cost-benefit relationships. As such, the analysis of this "yield gap" provides perspective for model evaluation. It is also essential for guiding sustainable intensification of agriculture worldwide to meet the increasing global food demand driven by population and income growth (van Ittersum et al., 2013).

Sensitivity analysis offers a tool applicable for addressing various aspects of model set-up and evaluation. Along with its applications in identifying parameters for model calibration, when based on observed ranges of parameter values, sensitivity analysis can also be used as a tool for model evaluation, for example helping to identify parameters that have the largest influence on the output. In the context of climate change impact assessment, by extending the analysis to cover climatic conditions outside previously observed ranges, sensitivity analysis can also be used to investigate impact model stability and behaviour across a range of plausible changes in input variables. Impact response surfaces (IRSs) provide one approach for analysing the sensitivity of specified impact variables to changes in key drivers (e.g. Fronzek, 2013; Rosenzweig et al., 1996; Ruane et al., 2014). These are discussed further in section 2.5.

1.4.3 UNCERTAINTIES IN CROP MODEL ESTIMATES

A common problem in model evaluation is that both the observations and simulations may be inaccurate for a number of reasons, leading to uncertainty about the main sources of error. Firstly, observations on actual crop yields and other crop components providing information of the real world that the simulations aim to represent, may be inaccurate or inappropriate for direct comparison against the simulated yields for a variety of reasons. Data are often made available as national statistics collected from a wide variety of sources (surveys, administrative sources, experts and various other data providers) and aggregated to the national level, such as yield data provided by EUROSTAT (EUROSTAT, 2014) or FAO (Food and Agriculture Organization of the United Nations, 2017). Thus, a single yearly yield value represents an average yield across a range of soil types, weather conditions, crop cultivars and varying management practices and occurrences of pests, weeds and diseases. Aggregation methods and quality of data may differ in different world regions complicating comparison between countries. On a regional level, datasets from official variety trials may exist providing good quality observations but the number of trials per region may be rather small and values are also often provided as aggregate values per region (Forkman et al., 2012). Individual study setups may exist where crop data is collected but the study setup or its emphasis may differ from the simulation setup at hand. In general, observations on crop characteristics other than yield are often not available, necessitating alternative approaches for approximating critical issues such as time of sowing, flowering and harvest and use of fertilizers and supplementary irrigation, needed for calibrating and running a process-based crop model.

The sources of uncertainty in crop modelling are typically classified to arise from: 1) input data, 2) model parameters and 3) model structure (Wallach et al., 2014; van Oijen and Ewert, 1999). Process-based crop models require a range of *input data,* such as data on weather, soil physical properties, crop/variety and crop management practices, that vary between sites and/or years and often can be measured. However, the values may be affected by errors in measurements and limitations in data availability at the appropriate spatial or temporal resolution, for example absence of a weather station at the site of interest (Wallach et al., 2014). While observations on temperature and precipitation may in many cases be available, daily data on variables such as solar radiation and vapour pressure may need to be derived from other variables. The exception is windspeed, that cannot be estimated from other variables and is often substituted by some average value. With respect to soil parameters, actual measurements are rarely available, but various approaches relying on e.g. expert opinion, predictive functions and GIS techniques may need to be used (Grassini et al., 2015).

In contrast to model input variables, values of *model parameters* can seldom be measured directly and often need to be estimated, where uncertainty can result from the chosen estimation technique, quality of the data set used or expert bias when relying on expert knowledge (Wallach et al., 2014). Additionally, model calibration may be inadequate or inappropriate and thus not capable of finding the best parameter values given the data (Seidel et al., 2018). *Model structural uncertainty* results from the fact that no model can comprehensively describe all relevant processes, including all explanatory variables of importance. On the other hand, several alternative approaches may be available for describing any given process where a chosen approach may turn out to not be the best option as new comparisons of different possible approaches are conducted (Seidel et al., 2018).

In climate change impact assessment additional uncertainties are associated with scenarios of future radiative forcing of the climate, with projections of the climate response to this forcing, and with the techniques used to downscale these projections to a finer spatial resolution (Giorgi, 2005; Olesen et al., 2007). Probabilistic projections of climate offer an approach for quantifying uncertainty in future climate by ascribing likelihoods to projections from large multi-model ensemble simulations for a given radiative forcing scenario. These attempt to sample aspects of climate model uncertainties such as initial conditions and parameter as well as structural uncertainties in the model design. Though largely exploratory (e.g. Harris et al., 2010; Knutti et al., 2005; Räisänen and Ruokolainen, 2006), they have also been formally adopted in climate risk assessment studies (e.g. Murphy et al., 2009). There is also uncertainty in how projected increases in [CO_2], implied by scenarios of future radiative forcing, will influence plant photosynthesis and water use. For instance, there are disparities in the magnitude of the CO_2 response in experimental results for crops grown in elevated levels of [CO_2] between studies using Free Air CO_2 Enrichment (FACE) and other methods using open-top chambers or greenhouses (Porter et al., 2014b).

Considerable progress has been made recently in quantifying and reducing uncertainty in crop model estimation in major international collaborative efforts such as the Modelling European Agriculture with Climate Change for Food Security (MACSUR) knowledge hub (www.macsur.eu) and the Agricultural Model Inter-comparison and Improvement Project (AgMIP;Rosenzweig et al., 2013). For example, several studies have estimated the uncertainty in crop yield estimates attributable to model structure for wheat (Asseng et al., 2013; Martre et al., 2015; Palosuo et al., 2011), barley (Rötter et al., 2012) and maize (Bassu et al., 2014) by comparing ensembles of widely used crop models, while another multi-model ensemble study focused on the improvement of wheat models through re-parameterisation and improved model descriptions (Maiorano et al., 2017). Working with ensembles of crop models as a means to address uncertainty is a rather recent development in the crop modelling community (see Rötter et al., 2011a), having parallels with climate modelling where methods of working with multi-model ensembles have been developed since the 1980s. As a result, valuable experience has been gained and some of the recommendations made are also applicable in the crop modelling context (Wallach et al., 2016).

2 MATERIAL AND METHODS

Figure 2 presents the arrangement of the set-up applied in this study, highlighting also the sections relevant for addressing uncertainties, sensitivities and options for adaptation. All components of the figure are described in the following sub-sections.

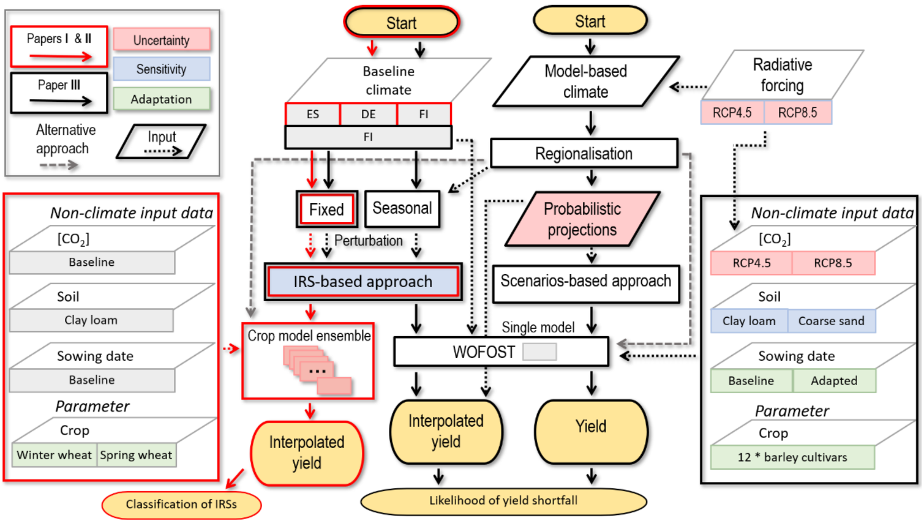

Figure 2 Arrangement of the study set-ups applied in papers **I** and **II** (red borders and arrows) and in paper **III** (black borders and arrows). Parts of the studies relevant for addressing issues associated with uncertainties and sensitivities, as well as for exploring adaptation, are highlighted respectively with pink, blue and green. Solid arrows indicate the flow of the stages of analysis and dotted arrows data providing input or processes informing various stages of analysis. Parallelograms are used to indicate the process of inputting data. Grey dashed lines indicate an alternative, commonly used, approach where regionalised model-based climate data are directly used as input into an impact model. Yellow symbols signify the start and the end results of the study-setups.

2.1 CROP MODELS

2.1.1 MODEL ENSEMBLE

An ensemble of 24 process-based wheat models run by 26 modelling groups was applied to study the sensitivities of wheat yields and the performance of crop models with perturbations of baseline temperature and precipitation (paper **I** and sub-sections 2.4.1 and 2.5.1) as well as to identify differences in the performance of individual crop models and possible reasons for such differences (paper **II**). In two cases the same model/version was calibrated separately and run by two modelling groups. In the ensemble, these were regarded as separate models. Paper **I** and its Supplement 1 provide an overview of all models, including key references and a characterisation of how they describe selected processes and treat environmental constraints. Most models were developed for the field scale, except for a few (i.e. CARAIB, LPJmL, LPJ-GUESS and MCWLA) that have been developed for regional assessment. While all models work on a daily time step, they apply a variety of approaches for describing processes that define, limit or reduce growth. This variation is largely due to different objectives driving the development of the models and is reflected in contrasting structure, model parameters and associated input data requirements.

As an attempt to summarise variations in the complexity of individual models in the ensemble, four key model processes were classified as being either detailed or simple (following a similar classification used by Asseng et al. (2013) for light interception). The selected processes are 1) light interception, 2) light utilisation and 3) yield formation, all describing aspects of crop growth, and 4) soil processes, encompassing the representation of the soil profile, water and nutrient dynamics. Results are summarised in Table 2 (criteria for the classification are described in Appendix 1).

Table 2. *Complexity of the 24 individual crop models applied in papers **I** and **II** for four key crop model processes; light interception and utilisation, yield formation and soil processes. Filled circles depict a more detailed process description; open circles a simpler representation. Numbers indicate the model ID and correspond with those reported in papers **I** and **II.** The ID and model name of WOFOST (25/26)**,** applied in paper **III,** are indicated in bold type. Shades of blue in the ID column define the overall complexity of an individual model: the darker the shade the greater the number of processes having a detailed description . The classification of complexity is based on the author's judgement using criteria reported Appendix 1.*

ID	MODEL NAME	LIGHT INTERCEPTION	LIGHT UTILIZATION	YIELD FORMATION	SOIL PROCESSES
1	AFRCWHEAT2	●	●	●	●
2	APSIM-Nwheat	○	○	●	●
3	APSIM-Wheat	○	○	●	●
4	AquaCrop	○	○	○	●
5	ARMOSA	●	●	●	●
6	CARAIB	●	●	○	○
7/8	CERES-wheat DSSAT v4.5	○	○	●	●
9	CERES-wheat DSSAT v4.6	○	○	●	●
10	CropSyst	○	○	○	●
11	DNDC	○	○	○	●
12	FASSET	●	○	○	●
13	HERMES	●	●	●	●
14	LINTUL-4	●	○	●	○
15	LPJ-GUESS	○	●	○	○
16	LPJmL	○	●	○	○
17	MCWLA-Wheat	○	●	○	●
18	MONICA	○	●	●	●
19	SALUS	○	○	●	●
20	SIMPLACE <Lintul2, Slim>	○	○	●	○
21	Sirius	●	○	●	●
22	SiriusQuality	●	○	●	●
23	SPACSYS	○	●	●	●
24	STICS	○	○	○	●
25/26	**WOFOST**	●	●	●	○

Based on the classification, most of the models in the ensemble have a detailed representation of soil processes (75%) and yield formation (62.5%), while processes associated with light interception and utilisation are more often described with relatively simpler approaches (Table 2). Three models out of the 24 simulate all four processes with detailed approaches (darkest shade of blue), but most models have an altogether simpler setup with one or two of the process descriptions being classified as detailed (two lightest shades of blue). For those models, the emphasis is most often placed on describing soil processes, which alone can still result in a fairly complex model setup.

2.1.2 THE WOFOST MODEL

One of the models included in the ensemble, WOFOST, version 7.1, was further applied to explore and illustrate the method for estimating the likelihoods of specified impacts occurring under a changing climate and assessing the effectiveness of adaptation options using spring barley as an example (paper

III). WOFOST is a dynamic, process-based, model that explains growth and production of annual field crops on the basis of underlying ecophysiological processes, such as photosynthesis, CO_2 assimilation and crop phenological development. Output variables such as attainable crop production, biomass and water use can be calculated with WOFOST given knowledge about weather conditions, crop and soil type and crop management factors (e.g. sowing date). The model follows the hierarchical distinction between potential and water- and nutrient-limited production. Factors reducing production such as pests, diseases and weeds are not considered in the model (Boogaard et al. 2011). With respect to the complexity of the four process descriptions described above (cf. Table 2), the soil processes in WOFOST are described on a relatively simple level, while the treatment is more detailed for CO_2 assimilation (light interception and utilisation) and yield formation.

The model was developed within the Wageningen "De Wit School" in the 1980s with original applications in the tropics. However, due to the general applicability of the biophysical core of the model, it has since been developed to be applicable across a wide range of conditions. In addition to its wide scientific use, it has also been applied for 25 years for operational crop yield forecasting in Europe within the Monitoring Agricultural Resources Unit (MARS) Crop Yield Forecasting System (MCYFS; Micale and Genovese, 2004) established by the Joint Research Centre (JRC), making it one of the longest running operational models. Its advantages also include the public availability of the full source code (de Wit et al., 2019).

2.2 STUDY SITES

Four study sites were chosen across a European transect to cover a range of climatic conditions for current cereal production (papers **I** and **II)**. The chosen sites represent contrasting environmental zones (Boreal, Continental, Atlantic Central and Mediterranean South) from northern to southern Europe (Figure 3). The aim was to choose sites where the choice of switching from winter wheat varieties to spring varieties, or vice versa, could be a possible adaptation option. The Finnish site, Jokioinen, represents predominantly temperature-limited conditions for cereal cultivation, whereas in Lleida (Spain) cultivation is mainly water-limited. To represent the climatically more favourable conditions of Central Europe, two sites in Germany were chosen, Dikopshof for simulations of winter wheat and Nossen for spring wheat. Both varieties were simulated for conditions at the Finnish and Spanish sites. The different conditions are reflected in the levels of observed yields over 1981 to 2010, which for wheat averaged 3300 kg ha^{-1} in Finland, 2500 kg ha^{-1} in Spain and 6700 kg ha^{-1} in Germany (FAOSTAT, 2018). For barley in Finland the equivalent average yield was 3200 kg ha^{-1}. In Finland spring varieties of wheat and barley are currently cultivated on a larger scale (Finnish Cereal Committee, 2014) whereas, in Germany and Spain winter wheat is the dominant type (Ministerio de Agricultura, 2018; Statistisches Bundesamt, 2018). In contrast to a focus on wheat in papers **I** and **II,** barley was chosen for paper **III,** which focused on Jokioinen**.**

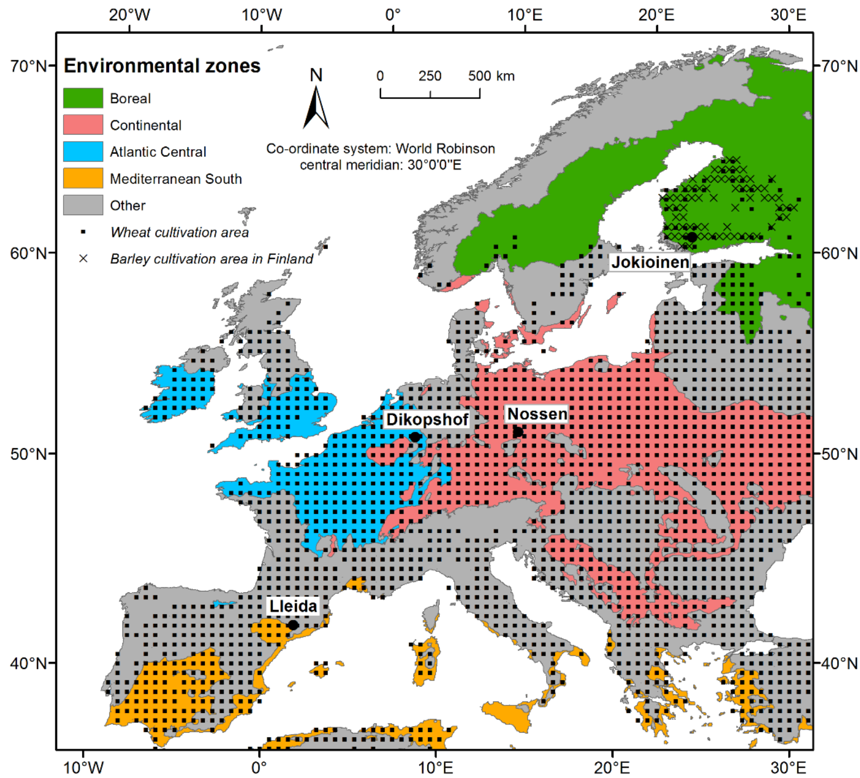

Figure 3 Locations of study sites superimposed on environmental zones as defined by Metzger et al. (2005). Black squares indicate the wheat cultivation area in Europe. Crosses are barley cultivation areas in Finland. Both datasets, adapted from Monfreda et al. (2008) and re-sampled to a 0.5-degree grid, depict land use circa the year 2000.

2.3 CROP AND SOIL DATA

Data for the four different sites, provided to the individual modelling groups for model calibration, consisted of phenological observations and yields for both crop varieties as well as management data on the time and depth of sowing, fertilisation, irrigation, tillage and residue management (papers **I** and **II**). Depending on the site and wheat variety, the data included observations from between 5 and 29 seasons. Modellers also had the option to use information either on the actual soil of the site or on a generalised soil type (clay loam) used across the sites in the model simulations. No specific guidance was given to the modellers as to the method of calibration applied, resulting in a variety of approaches being selected ranging from no calibration to application of Bayesian methods. For paper **III**, the calibration of WOFOST was based on a previous calibration of the model applying a manual approach and utilising data from Finnish variety trials and field experiments that included phenology and yield (Rötter et al., 2011b).

For the actual model simulations of papers **I** and **II,** for the Finnish site and the two German sites, yearly sowing dates were determined for each of the baseline years (1981–2010) based on observations and applied with all perturbations of the baseline climate (conducted as part of the sensitivity analysis of the impact models, see sub-section 2.5.1). Due to the absence of observed sowing dates for Spain, one fixed sowing date (day of the year [DOY] 302) was identified based on local expertise and applied for all years and both spring and winter wheat. The assumption was that spring wheat could safely overwinter in the relatively mild winters so that the same sowing date could be applied for both wheat varieties (Mínguez et al., 2007). For paper **III**, sowing dates, both for the baseline and perturbed temperatures, were determined using a temperature-based calculation method, developed in the study. The method is based on the relationship between temperature and observed sowing dates for barley across 20 regions in Finland, from 1988 to 2012.

Simulations were allowed to continue in papers **I** and **II** without a defined end date and in paper **III** to the end of the year. However, in post-processing of simulation results a harvest cut-off was applied to avoid the unrealistic outcome of model simulations carrying on past plausible windows for harvest (i.e. extending to the end of the year or even into the following growing season). In papers **I** and **II** a fixed cut-off date, based on expert judgement was applied (DOY 258 for Finland and Spain and DOY 274 for the German sites). In paper **III** the cut-off was defined dynamically for each simulation separately following a temperature-based method approximating the occurrence of the first frost. During the baseline period the average cut-off was DOY 267 (end of September). If a maturity date exceeding the harvest cut-off was reported in the results, grain yield and nitrogen content were set to zero and all other variables assigned missing values.

Clay loam was used as the default soil type at all sites and in all three studies. Additionally, in paper **III** model simulations were also performed for coarse sand as a test of model sensitivity to less favourable soil conditions with respect to water holding capacity. For comparison of simulated against observed yields in papers **I** and **II**, regional grain yield observations for the baseline period (1981–2010), where available, were obtained from FAO for Finland (FAOSTAT, 2018), Eurostat for eastern (Nossen) and western (Dikopshof) Germany (EUROSTAT, 2014) and the Spanish Ministry of Agriculture for northern Spain (MAGRAMA, 2010) and previous statistical yearbooks. In paper **III** yield observations for barley were obtained from Finnish variety trials at Jokioinen for the entire baseline period (Kangas et al., 2010) and presented as aggregated annual yields for a cultivar group classified as having an intermediate development rate based on the length of the growth cycle (Palosuo et al., 2015).

2.4 CLIMATE DATA

2.4.1 OBSERVED CLIMATE

Observed daily weather data were obtained from weather stations at the four study sites for the baseline period of 1981 to 2010. In addition, data were collected for the preceding year 1980 to cover the start of the growing season of the first harvest (1981) of winter wheat (papers **I** and **II**). The data set consisted of minimum and maximum temperature, precipitation, global solar radiation, wind speed and various measures of humidity to account for the different data requirements of individual models. Wind speed for Jokioinen, recorded at 10 m height, was converted to 2 m height assuming a logarithmic wind profile following Allen et al. (1998; their eq. 47). Details on the procedures for deriving values for missing variables as well as the sources of the data are described in Supplement 2 of paper **I**.

2.4.2 FUTURE PROJECTIONS

In paper **III** future climate change, required for estimating impact likelihoods, was represented as regional temperature and precipitation changes, relative to 1981–2010, as projected by a set of probabilistic projections of temperature and precipitation changes. The changes are based on projections from the CMIP5 ensemble of GCM simulations and relate the changes to different levels of radiative forcing described by the RCPs (see Paper **III**, sub-section 2.3.3 and Supplement 1). In our study we focused on changes, corresponding with RCP4.5 (intermediate) and RCP8.5 (high) for seven future 30-year time periods (2011–2040, 2021–2050, ..., 2071–2099). A resampling method for deriving probabilistic projections, developed by Räisänen and Ruokolainen (2006), was applied that accounts for natural climate variability and structural uncertainty across the CMIP5 ensemble of GCMs. In the approach, the sample size of an ensemble is significantly increased from the number of GCM simulations. This is done by first identifying the global mean temperature change of the target future time period (e.g. 2021–2050), then identifying any other two time periods within the simulations with the same global mean temperature change between the two. The resampled changes for the future time period in question (e.g. 2021–2050) are then the regional changes in temperature for that global mean change. The resampled precipitation changes are those that match the identified regional temperature changes.

The approach lacks some of the aspects of uncertainty associated with an earlier set of probabilistic projections of changes for medium emissions (SRES AIB) during the 21st century based on a perturbed physics experiment using model projections from the CMIP3 archive (Harris et al., 2010), but being based on CMIP5 it provides more up-to-date projections. The SRES-based projections were applied in papers **I** and **II** to provide information on the projected ranges of temperature and precipitation changes for the sensitivity analysis along with supplementary information from the CMIP5 ensemble of GCMs.

The GCM-based projections of temperature and precipitation changes include seasonal variation throughout the year for RCP8.5 for the last time period (2071–2099) for the individual (resampled) model projections as well as the ensemble mean (Figure 4, left panels) and across the different time periods for the ensemble mean (Figure 4, right panels). For the individual model projections of precipitation changes there is a lot of variability in the seasonal pattern, but otherwise the signal shows a very consistent seasonal pattern for temperature and for precipitation across the different time periods. In the context of paper **III** this seasonal pattern was accounted for in the sensitivity analysis of crop responses to changes in temperature and precipitation (for details, see Supplement 4, equations S4–S7 of paper **III)**, while in papers **I** and **II** the changes were applied as constant annual changes.

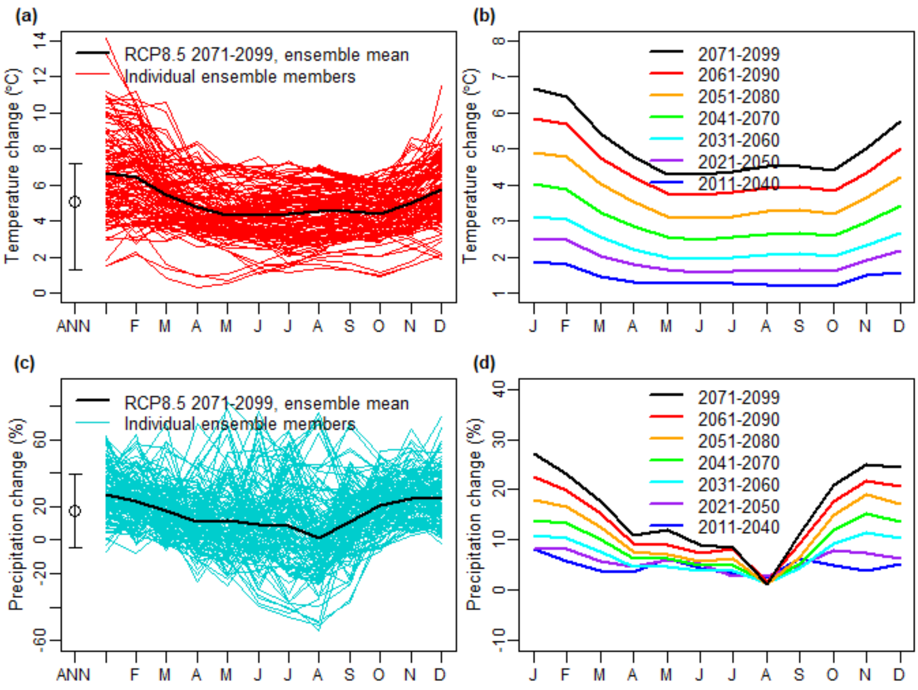

Figure 4 Projected changes for RCP8.5 at Jokioinen Finland for the latest 30-year period (2071–2099) for the ensemble mean (black line) and individual (resampled) members of the CMIP5 ensemble for temperature (a, red lines) and precipitation (c, turquoise lines) throughout the year. Individual ensemble members cannot be distinguished, and the lines illustrate the overall variation within the ensemble including both the original and resampled members. The variation across the model ensemble for the mean annual changes is shown as whiskers (ANN) with the ensemble mean depicted with a circle. Panels on the right depict the monthly pattern of the ensemble mean changes (RCP8.5) for seven future time periods for temperature (b) and precipitation (d). Note that the vertical scales between (a) and (b) as well as between (c) and (d) are different.

2.5 IMPACT RESPONSE SURFACES AND THEIR APPLICATIONS

2.5.1 CONSTRUCTION OF IMPACT RESPONSE SURFACES

The method used as a basis for most analysis in this study was the construction of two-dimensional impact response surfaces (IRSs) depicting the sensitivity of modelled yield across a large range of projected changes in mean annual temperature and precipitation (for an example for Jokioinen see the range of projected annual changes in Figures 4a and c depicted by whiskers). The IRSs are constructed by plotting modelled yield results as contour lines by bi-linear interpolation at each increment of temperature and precipitation changes, thus providing a visual interpretation of modelled crop behaviour under climate change (see coloured surfaces in Figure 5).

In the construction of IRSs baseline values of observed daily temperature and precipitation were perturbed using a simple "change-factor" approach (Diaz-Nieto and Wilby, 2005). In papers **I** and **II** mean annual changes were applied systematically as a constant change throughout the year and values of all other variables, including [CO_2], were kept unchanged from their baseline values so that the focus of analysis could be on the pattern of yield response to changes in climate across different crop models. In paper **III** the focus was on estimating likelihoods of crop impacts, thus placing a need for adding more realism to the model simulations. This was implemented firstly through seasonal weighting of the annual changes, accounting for the seasonal pattern of future changes projected by climate models (see Figure 4). Consequently, for a given annual change there are seasonal variations throughout the year, e.g. +2°C translates into a change of +1.7°C during the summer and +2.5°C in the winter. Secondly, relative humidity was assumed to remain unchanged from the baseline with changes in temperature (Lorenz and DeWeaver, 2007), requiring daily vapour pressure to be adjusted as a function of temperature (Allen et al., 1998). Thirdly, effects of projected changes in the level of [CO_2] were accounted for through adjustments of crop growth parameters reflecting CO_2-related changes in plant response (for details, see Paper **III**, Supplement 3, equations S1–S3).

Individual IRSs were created for a unique combination of parameters (baseline year, crop, site, soil type, [CO_2], sowing method, cultivar...) demonstrating differences in modelled crop response across baseline harvest years and in paper **III** additionally across soil type, future [CO_2] and different adaptation options. The options for adaptation were demonstrated through alternative sowing times and barley cultivars. The effect of the time of sowing was assessed by comparing situations of non-adapted sowing where baseline sowing dates are fixed for all perturbed climates (hereinafter referred to as baseline sowing), with situations of autonomous adaptation, represented through "optimal" sowing with respect to temperature (referred to as adapted sowing – see section 2.3). The different barley cultivars represented existing cultivars (Scarlett and Annabell) as well as ten versions of designed cultivars defined simply on the basis of their thermal requirements (°C d) from emergence to anthesis (TSUM1) and from anthesis to maturity (TSUM2), considered as providing potential for the crop to exploit more effectively the changing temperature conditions during the growing season. Scarlett was used as the default cultivar for the simulations with barley.

2.5.2 MEASURES FOR ANALYSING MODEL ENSEMBLE RESULTS

For analysing results produced by the ensemble of 26 wheat models covering the baseline period (1981–2010) at the different study sites across Europe, a variety of measures were plotted as IRSs, describing aspects of average yield responses and their dispersion (paper **I**). Across models, ensemble medians were used to analyse patterns of average (30-year mean) yield responses, depicted as changes relative to the baseline, and inter-quartile range (IQR; from the 25th to the 75th percentile) for the spread of the response. Across baseline years, to address aspects of inter-annual variability associated with the model ensemble median responses, two alternative measures were used, a measure of year-to-year reliability (percentage of years having a yield exceeding a defined threshold yield) focusing on the low end yield responses (10th percentile), and the coefficient of variation (CV; the ratio of standard deviation to mean) across the 30 years, accounting for the full distribution.

2.5.3 CLASSIFICATION OF CROP MODEL RESPONSES

Two alternative approaches were developed for classifying crop model responses plotted as IRSs: the expert diagnostic approach (EDA), relying on an understanding of the yield responses represented as IRSs, and a statistical diagnostic approach (SDA), based on the comparison of the pattern and magnitude of response across the IRS, without attempting to interpret these features (paper **II**). In the EDA, the two defining characteristics of the response were the location of the maximum yield compared to the baseline with respect to change in both temperature and precipitation (black box in each IRS of Figure 5), and the strength of response, defined as the rate of change in yield expressed separately with respect to temperature and precipitation relative to the location of the maximum (illustrated with arrows on the right hand IRS in Figure 5). The SDA applies a hierarchical clustering method using a distance metric (*d*) that combines the spatial correlation and Euclidian distance between IRS pairs. For details of the classification approaches see paper **II** and its Supplement, equations 1–3**.** These two approaches were further used to explore whether particular properties of models and the modelling set-up (site, genealogy, calibration method used and specific model process descriptions) could be related to different IRS patterns of modelled yield response to temperature and precipitation changes.

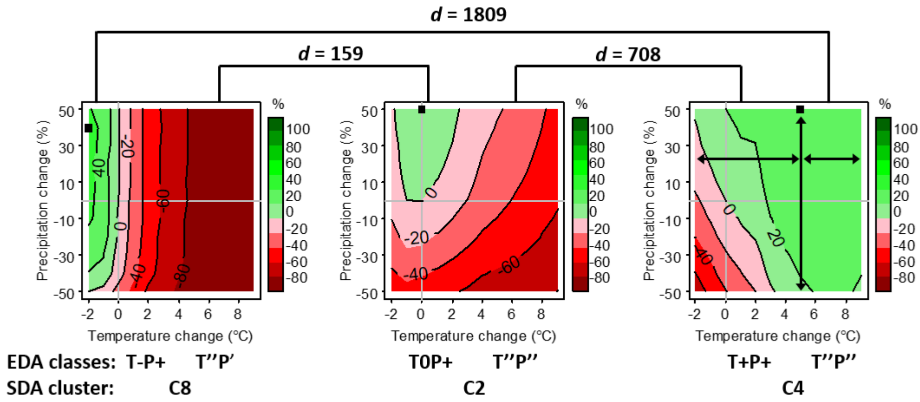

Figure 5 Examples of IRSs representing three different classes of IRSs as well as different clusters as defined by the expert diagnostic approach (EDA) and the statistical diagnostic approach (SDA), respectively. The EDA classification is based on the location of the maximum yield (small black box on each IRS) and the rate of change in the yield response identified separately for both sides of the location of the maximum with respect to both temperature and precipitation (illustrated with arrows in the IRS at the right and see Table 3). SDA clustering is based on a distance metric *d* where a smaller number indicates that the IRSs are closer to each other in the shape and magnitude of the pattern in contrast to those with a higher value.

2.5.4 ESTIMATION OF IMPACT LIKELIHOODS

In this study, the approach applied for estimating risks and possible benefits of climate change on crop production involved making use of projections of climate change interpreted probabilistically, defined as joint frequency distributions of temperature and precipitation change (paper **III**). These projections, representing different time periods in the future and conditional on a specific RCP, were then superimposed on top of IRSs depicting the response in yield to changes in the same two drivers. The approach was then applied to estimate the likelihood of a critical impact occurring.

The likelihood of barley yield shortfall (representing a critical impact to barley production) was evaluated with respect to a specified threshold yield by integrating over the parts of the joint frequency distribution of future climate falling below the defined threshold. In practice this involved computing the percentage of re-sampled projections lying on top of regions of the IRS with yields lower than the threshold yield (Figure 6a). This process was repeated for individual future time periods where both the IRS and the probabilistic projection of future climate are representing the same time period, with the effect of [CO_2] accounted for. Consequently, the evolution of the likelihood of yield shortfall throughout the century could be presented. The approach was also applied to assess the effect of associated uncertainties on average yield estimates due to alternative projections of future climate change and inter-annual variability, but is not described here (see Paper **III,** Figure 6).

As a means of evaluating the IRS-based approach of estimating impact likelihoods, the likelihood of barley yield shortfall was also estimated using a more conventional scenario-based approach. Here, crop model simulations were performed for each year of the baseline period and then for the same 30 years perturbed according to the seasonal changes indicated by the ensemble of resampled climate projections. For the analysis of likelihoods, the yields were then averaged to 30-year mean yields for each resampled climate projection. The likelihood of yield shortfall was then calculated as the number of projections with 30-year mean yields lying below the threshold relative to the full ensemble of resampled climate projections (Figure 6b).

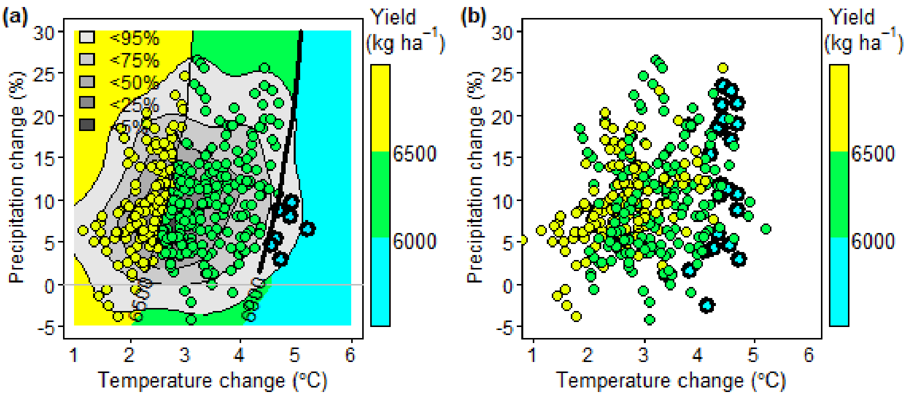

Figure 6 Methods for evaluating impact likelihoods under projected climate change. Projections: circles are the resampled GCM projections of future period-mean temperature and precipitation change relative to 1981–2010. Grey shading depicts relative frequencies fitted to these projections. Methods: (a) IRS-based approach – yields from the underlying IRS (coloured surface) are interpolated to the circles, hence colours of circles match the coloured IRS background. (b) Scenario-based approach – colour of circles depicts the level of yield estimated by directly applying the projected changes to baseline weather. In both (a) and (b) circles with a thick black outline identify projections giving yields below a threshold of 6000 kg ha$^{-1.}$ Likelihood of yield shortfall is calculated as the ratio of the number of circles giving yields below and above the threshold. For details, see text.

3 RESULTS

3.1 MODEL PERFORMANCE UNDER BASELINE CONDITIONS

Simulated and observed yields for Jokioinen, Finland during 1981–2010 have been examined to get an impression of yield variations and model performance under baseline conditions for both wheat and barley (Figure 7). Examination of the yield series as absolute values (kg ha $^{-1}$) enables levels of yields to be assessed with respect to different sources of data (Figure 7a), while expressing the yields as normalised anomalies relative to the long-term mean gives a more standardised impression of how yields respond to annual weather (Figure 7b).

Observations covering the baseline period were available from FAO as national statistics for both wheat and barley. Comparison of these yield series showed the yearly yields to be very similar throughout the baseline (not shown). Consequently, only wheat observations are shown. For barley, yields from Finnish official variety trials at Jokioinen were examined. These two sets of observations (FAO national statistics and variety trial data) differ from each other quite considerably. Yields from the variety trials exceed the national statistics in nearly every year; 30-year means are 5299 kg ha^{-1} and 3306 kg ha^{-1} for variety trials and national statistics, respectively. The inter-annual variation is more pronounced in the variety trial data (CV = 26%) as opposed to the national statistics (CV = 17%). There are also some differences in the year-to-year pattern, though both data sets reflect low yields in years with particularly adverse conditions for agriculture such as 1987 and 1999. Note that the site wheat data provided for model calibration in paper I (also shown in Figure 7) shows a yet different kind of pattern from year to year.

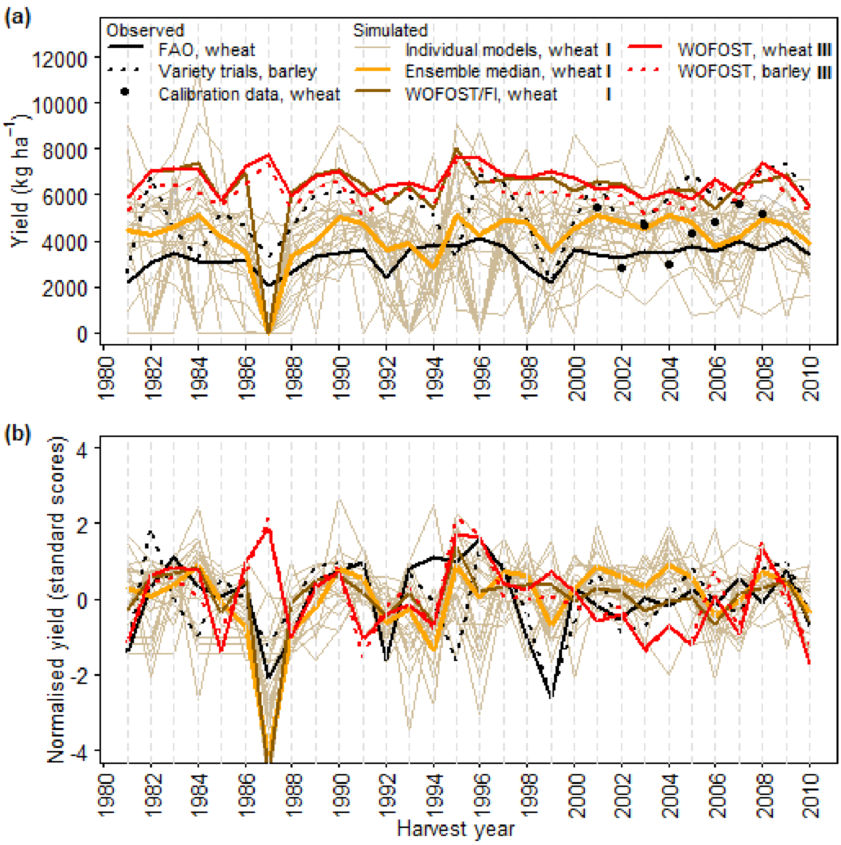

Figure 7 Mean annual dry matter grain yields (kg ha^{-1}) of spring wheat and barley for the 1981–2010 baseline period at Jokioinen, Finland: (a) absolute and (b) normalised (standard scores, i.e. standard deviations relative to the 1981–2010 mean). Observed yields are presented in black: solid line shows national yield statistics for wheat, dashed line data from official variety trials for barley at Jokioinen, Finland and dots Jokioinen site data provided for calibration in paper **I**. In (b) the data for the black lines is adjusted to account for long-term trends assumed to be unrelated to weather by removing a linear trend. Tones of brown and orange depict simulation results from paper **I**. Beige lines: yields simulated by individual models. Dark brown: WOFOST/FI (ID 25) simulation results highlighted. Orange line: multi-model ensemble median yields. Red lines depict simulations results using the WOFOST modelling set-up of paper **III**. Dashed line: simulation results for barley as presented in the paper. Solid line: simulation results for wheat using the wheat calibration file from paper **I** and the modelling set-up of paper **III**.

The simulation results of the multi-model ensemble for wheat from paper **I** (beige lines, Figure 7) vary widely between models. Here, we show only results for spring wheat, but the findings for winter wheat are largely similar. In the model ensemble, individual models can be found to simulate yields that are either considerably higher or substantially lower than the ensemble median (orange line). Of this set of individual model results, the yields as simulated by the WOFOST/FI model (model ID 25) are highlighted for comparison against other simulation results by the same model. Comparison is made against barley and wheat yields simulated following the modelling set-up designed for paper **III**, where e.g. the sowing date and cut-off date at the end of the growing season are defined differently from those in paper **I**. The calibration file for wheat is the same in all spring wheat simulations. These three sets of WOFOST simulations give fairly similar results throughout the baseline period, apart from a big difference in 1987 for which the WOFOST/FI results (dark brown line) give zero yield and the two simulations following the paper **III** set-up result in the highest yield of the time series. In the WOFOST/FI simulations the crop does not reach maturity before the cut-off and thus the yield is set to zero, whereas, in the simulations for barley and wheat following the paper **III** set-up the crops are sown early resulting into a long growing season reaching maturity before the cut-off.

In comparison to simulation results by the other models and to the multi-model ensemble median, yearly yield levels simulated by WOFOST are at the top end of the range, but the year-to-year pattern is often fairly similar to the ensemble median. Of the years where WOFOST results deviated most from the other simulation results, 1996 and 1999 were both years with drought effects on crop production. The difference between the simulations for spring wheat and barley are mainly that wheat yields are often higher than those for barley.

In comparison to observed data, no statistical measures of correspondence between observed yields and those simulated by the individual models and the multi-model ensemble median are provided due to issues in the observed datasets complicating direct comparison (see discussion in sub-section 4.1). On visual inspection some large deviances during individual years can be found. With respect to the yield levels, a large gap between actual farmers' yields and simulated yields can be found when comparing the highlighted individual model results against the national yield statistics by FAO. On the other hand, in comparison to the official variety trial data the gap is much smaller during many years and during some years the observed yields are higher than those simulated. In comparison to the individual model results, the model ensemble median produces yields closer to the national statistics, with the gap between the two being on average 916 kg ha^{-1} for spring wheat during the baseline period (2033 kg ha^{-1} for winter wheat).

Results for the baseline for Germany and Spain, for both spring and winter wheat, are presented in paper **I** (Figure 3 and Figure S1). In Lleida, Spain the spread across simulated yields by individual models was the greatest, as well as the gap between the ensemble median and the observed yields (on average the yield gap for spring wheat was 2706 kg ha^{-1} and for winter wheat 1603 kg ha^{-1}). In contrast, at the German sites the ensemble median and observed yields were on average much closer to each other (yield gap for spring wheat was 233 kg ha^{-1} and for winter wheat 415 kg ha^{-1}) and the year-to-year pattern was simulated rather well.

3.2 SENSITIVITY OF CEREAL YIELDS TO TEMPERATURE AND PRECIPITATION IN EUROPE

Mean yields of both wheat and barley were found to be sensitive to changes in temperature, while assuming baseline precipitation, declining with greater warming at all examined sites across Europe (Figure 8a). Again, key results for winter wheat are in line with those for spring wheat, though not shown here. The mean yields for barley in Finland as simulated by WOFOST, although also declining with warming (~3% average yield decline per 1°C between 0 and +8°C), were found to be less sensitive to temperature in comparison to the response of spring wheat as simulated by the model ensemble (~-5% per 1°C), as well as by WOFOST as part of the ensemble (~-7% per 1°C).

With cooling all simulated results for Finland show a steep decline in yields, whereas for Germany the yields increase and for Spain stay approximately at the baseline level. In the simulations for wheat, temperature and precipitation adjustments were applied evenly throughout the year. For barley, the effect of applying a seasonal weighting to the changes (cf. sub-section 2.4.2) shows as a slightly reduced yield decline with warming in comparison to the simulations with a constant annual change (~-3% average yield decline per 1°C between 0 and +8°C with seasonal weighting and ~-4% with constant annual change). With respect, to the sensitivity of yield results to different levels of [CO_2], the higher the level of assumed [CO_2] the higher the yields are, though the pattern of yield sensitivity to temperature and precipitation change at higher [CO_2] resembles that under 360 ppm (Figure 8). The inter-model spread is the highest for Spain and lowest for Germany especially at the warmest temperatures.

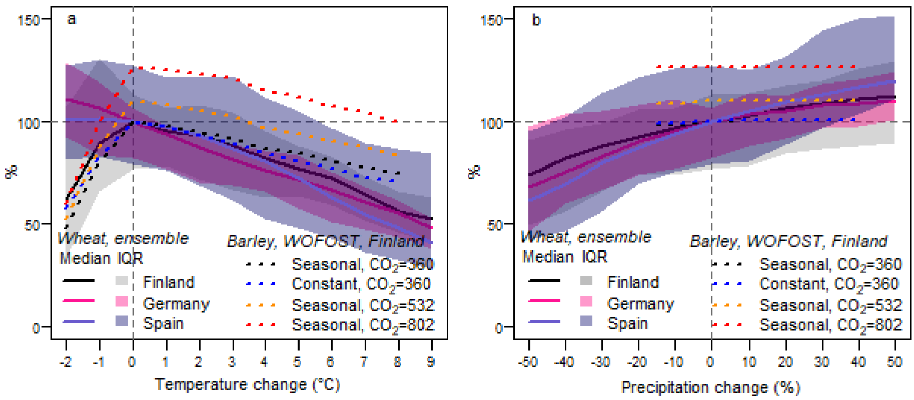

Figure 8 Ensemble median response (solid lines) and inter-quartile range (IQR – coloured bands) of period mean dry matter spring wheat yield (%) relative to the baseline (1981–2010) climate across 24 crop models at Jokioinen, Finland, and 25 models at Nossen, Germany and Lleida, Spain at 360 ppm. Dashed lines are for barley simulated with WOFOST for Jokioinen, Finland by applying seasonal weighting to changes to the baseline climate at different levels of [CO_2]: 360 ppm (black), 532 ppm (the 2071–2099 value for RCP45 – orange) and 802 ppm (the 2071–2099 value for RCP85 – red). Blue dashed lines are for a constant annual change to the baseline climate at 360 ppm. Yields are presented for changes in (a) temperature with baseline precipitation, and (b) precipitation with baseline temperature. Baseline values are scaled to 100%.

For changes in precipitation, while assuming baseline temperature, a positive relationship between simulated wheat yields and precipitation was found across the sites with the effect being stronger the further south the location is in the transect. The association of yield changes with precipitation was less linear than with temperature, with greater sensitivity to reduced than increased precipitation. For example, in Lleida, Spain, where the sensitivity was the greatest the yield decline was -10% per 10% change in precipitation under reduced and +7% per 10% for increased precipitation (Figure 8b). Conversely, the yields for barley in Finland as simulated by WOFOST were found to be almost completely insensitive to precipitation at baseline temperature, with the method of applying the changes to the baseline climate (seasonal vs constant) having virtually no effect on the yield (note, however, that the range of decreases to precipitation extends only to -15% for barley as opposed to -50% for wheat). Similarly, as with temperature, higher [CO_2] results in higher yields but the sensitivity to changes in precipitation is unchanged. The inter-model spread is again highest for Spain across all changes in precipitation and lowest for Germany for the biggest changes.

By expressing yields at different temperature changes relative to the baseline yield of each respective dataset, i.e. alternative ways of applying the changes to temperature (seasonal weighting vs constant annual change) and assuming different levels of [CO_2] concentration, it was found for barley that the higher the level of [CO_2] the more it compensates for the negative impacts of higher temperature, with the impact strongest at the biggest temperature changes (Figure 9). The effect of applying a seasonal pattern rather than constant change to baseline temperature adjustments also evens out the decline in yields with warming. For cooling, the only difference can be seen between the seasonally weighted and constant annual changes, with the constant change method resulting in slightly less decline. The level of [CO_2] does not affect the yield response with cooling.

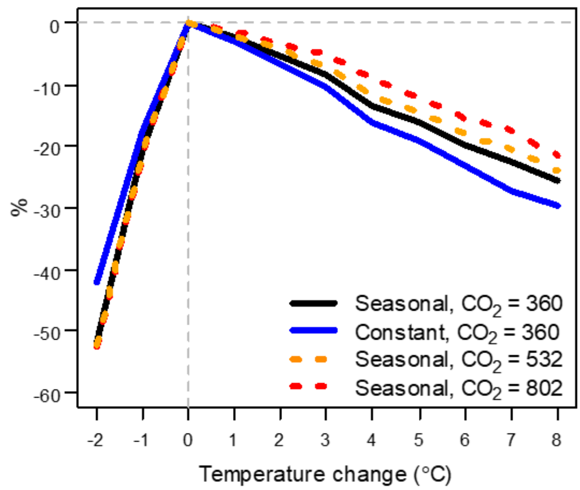

Figure 9 Period mean dry matter barley yield (%) relative to the baseline (1981–2010) climate at Jokioinen, Finland simulated by WOFOST for changes in temperature at baseline precipitation. Solid lines assume 360 ppm [CO_2] with seasonal weighting applied to the changes to temperature (black line) or by applying a constant annual change (blue line). Dashed lines assume elevated [CO_2]: 532 ppm (the 2071–2099 value for RCP45 – orange) and 802 ppm (the 2071–2099 value for RCP85 – red). The period mean yield of each set of simulations at zero temperature change (baseline climate) is scaled to zero. Yields at different temperature changes are expressed relative to each respective baseline yield.

For examining the interactions of temperature and precipitation and consequent effects on yield, simulation results for barley and wheat were plotted as the percentage change in 30-year mean yields relative to the baseline constructed for each crop variety and for wheat for the different sites as two-dimensional IRSs. For wheat the results are analysed as model ensemble median responses over the full set of models applied in paper I (24 models for Finland and 25 for Germany and Spain). The barley response is based on one model, WOFOST. Results for spring barley and wheat are shown in Figures 10 a–d.

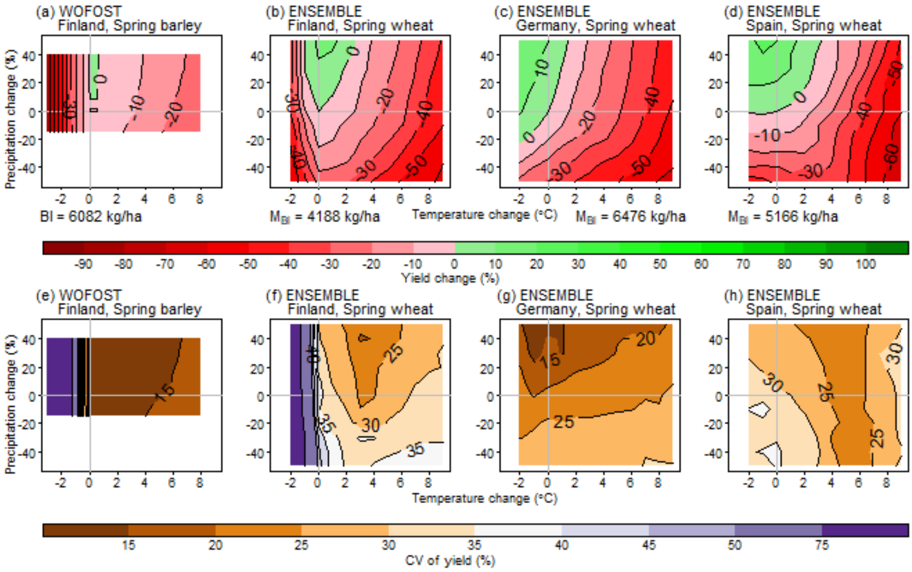

Figure 10 Percentage changes in thirty-year mean dry matter grain yields relative to the baseline (Bl; 1981–2010) climate (top row) and coefficient of variation (CV) of annual yields (bottom row) for changes in temperature (x-axis) and precipitation (y-axis) relative to the baseline climate. a) and e): spring barley at Jokioinen, Finland simulated by WOFOST, following the modelling setup of paper **III**; (b–d) and (f–h): spring wheat for the ensemble median (M) of 24 crop models at Jokioinen, Finland (b and f) and 25 models at Nossen, Germany (c and g) and Lleida, Spain (d and h), following the modelling setup of paper **I**. Note, that the range of changes to temperature and precipitation differed in papers **I** and **III** resulting in blank areas over parts of plots (a) and (e) with no simulation results.

With spring wheat, the range of precipitation changes showing as drying emphasise the yield declines found with warming, particularly in Spain (Figures 10 b–d). When comparing the patterns of responses of barley and wheat in Finland (Figures 10 a & b) in the equivalent IRS space, a similarity can be seen in the shape of response, although the barley response depicts lesser sensitivity to warming. Although the comparison is made between the results of an individual model, WOFOST, and those of a multi-model ensemble, it seems fair to assume that greater drying would cause a similar decline in yields for barley as with wheat. Further, the plots communicate clearly the location of maximum yields.

Classification of the yield patterns depicted on the IRSs (in terms of absolute yields kg ha^{-1}) using the EDA confirms that both spring barley and wheat in Finland have a relatively stronger response to temperature than precipitation, while in Spain, both types of wheat have stronger response to precipitation than temperature (Table 3). This applies also to winter wheat at Dikopshof, Germany, while the response for spring wheat in Nossen, Germany and winter wheat in Jokioinen, Finland have a more mixed response to both drivers. The maximum yields are found with all examined cereals at all sites with increases to precipitation. Apart from Germany, where maximum yields are found under the greatest simulated cooling, maximum yields are located within ±1 °C of baseline temperatures (see also Figures 10 a–d.

Table 3. *Classification of IRS patterns of 30-year mean yields according to the expert diagnostic approach (EDA) for spring barley (simulated by WOFOST, with the modelling setup of paper **III**) and spring and winter wheat (multi-model ensemble median from papers **I** and **II**) at sites in Finland, Germany and Spain. Colour-codes are based on the strength of the yield response being either weak (') or strong ("); the location of maximum yield relative to the baseline is defined as: within ±1 °C or 10% of the baseline (0), for increases of > 1 °C or >10% (+), or equivalent decreases (-), respectively for temperature (T) and precipitation (P).*

Crop	Finland	Germany	Spain
Spring barley			
Spring wheat			
Winter wheat			

KEY:

Strength of response		Location of maximum yield relative to baseline	
T'P'		T0P+	
T'P"		T-P+	
T"P'			

By plotting the coefficient of variation (CV) across the 30 years, which accounts for the full distribution of yield responses, aspects of inter-annual variability in cereal yields can be analysed (Figures 10 e-h). Again, for both varieties of wheat this is expressed as the model ensemble median response. At the Finnish and German sites, similarities can be seen in the shape of the CV and the median response. The CV increases primarily in similar regions of the plot as where median yields are declining. However, for Germany there is a contrast in the patterns with respect to the dominant variable, which for the CV is clearly precipitation-dependent while the median yields are strongly temperature-dependent. The reverse applies at Lleida, where temperature dominates the inter-annual variability while median yields are strongly influenced by precipitation. With winter wheat, the biggest difference to the results for spring wheat is the pattern of the CV response at the Spanish site, where inter-annual variability is almost exclusively affected by precipitation, with yields being more variable with drying. When comparing the patterns of CV responses of barley (as simulated by WOFOST) and wheat (multi-model ensemble median) in Finland, it can be clearly seen that there is much less variation in the year-to-year yields of barley than of wheat, although there were similarities in the mean yield responses of the two.

Patterns of wheat yield reliability, defined as the percentage of years when yield is above the 10th percentile of the baseline yield, appeared to track those of the changes in mean yields relative to the baseline (see Figure 7 for spring wheat and Figure S3 for winter wheat in paper **I**). At all sites, reliability declined with increasing temperature and decreasing precipitation, but the response was slightly shifted along both axes and the rates of decline in reliability differed between locations. For example, in Germany and Spain the highest reliability was achieved with a slight decrease to temperature and in Finland with a slight increase.

3.3 IDENTIFICATION AND INTERPRETATION OF DIFFERENCES IN SIMULATED YIELD RESPONSES

In papers I and II, the IRSs of yield changes relative to the baseline at each site are constructed from the median responses of a large model ensemble. In addition to examining this consensus view of how different models simulate the joint effects of temperature and precipitation, it is also important to study the behaviour of individual models contributing to the ensemble response. Figure 11 presents examples of IRSs for spring wheat at the Finnish site, demonstrating the variation in the location of the maximum yield with respect to temperature and in the rate of yield change when moving across the IRS along the x- and y-axis. The full set of IRSs for the different sites is presented in Paper **II**, Supplement 2. Clearly, individual model responses can differ quite considerably from the ensemble median response shown in Figure 10b.

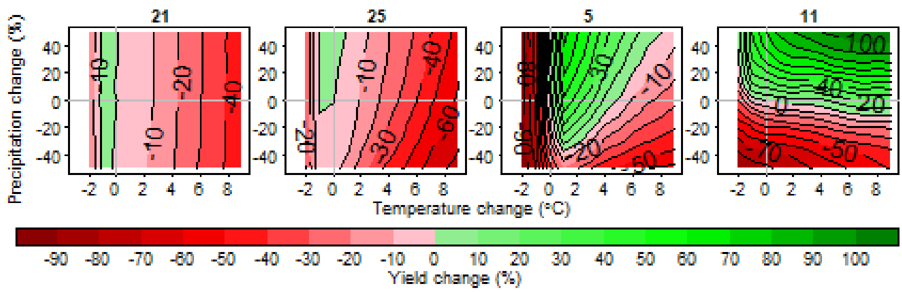

Figure 11 Thirty-year mean changes in spring wheat dry matter grain yields simulated by individual models of the multi-model ensemble for changes in temperature (x-axis) and precipitation (y-axis) relative to the baseline climate at Jokioinen, Finland. The number above each small plot is the model identification number (see Table 1 in Paper **I**).

Inter-model variability was investigated across all sites using the IQR (see Figure 5 in paper **I**). Model responses of spring wheat diverge most in Finland for cooling across the range of precipitation changes. In Germany, the inter-model variability is greatest with combinations of drying and cooling, while in Spain the IQR is greatest with warming combined with drying.

In paper **II**, the individual IRSs of different models at different sites for spring and winter wheat were classified according to two different approaches, the EDA and the SDA, to allow analysis of the varying patterns in the model ensemble. In the SDA, eight clusters were identified independently for each crop, with varying numbers of cluster members in each. Consequently, the patterns averaged across each cluster cannot be compared with each other between the two crops. The clustering method, based on a statistical algorithm measuring pattern similarity, allocates low cluster numbers to dominant patterns having the highest frequency of cluster members. Although not comparable quantitatively between the crops, the patterns in the two largest clusters (C1 and C2) show close correspondence qualitatively between spring and winter wheat (Figure 12). In both clusters the yield maximum is found with some cooling and increase to baseline precipitation. In C1 the temperature response is much stronger compared to precipitation, while in C2 the response is more mixed. For C3, which has fewer cluster members, the maximum yield is found with warming, though with greater warming for winter wheat than for spring wheat. Both crops show a strong negative response to cooling. Clusters 4–8 each contain 1 to 4 members, which exhibit patterns of behaviour that differ from the majority of ensemble members (see Figure 5 in Paper **II**).

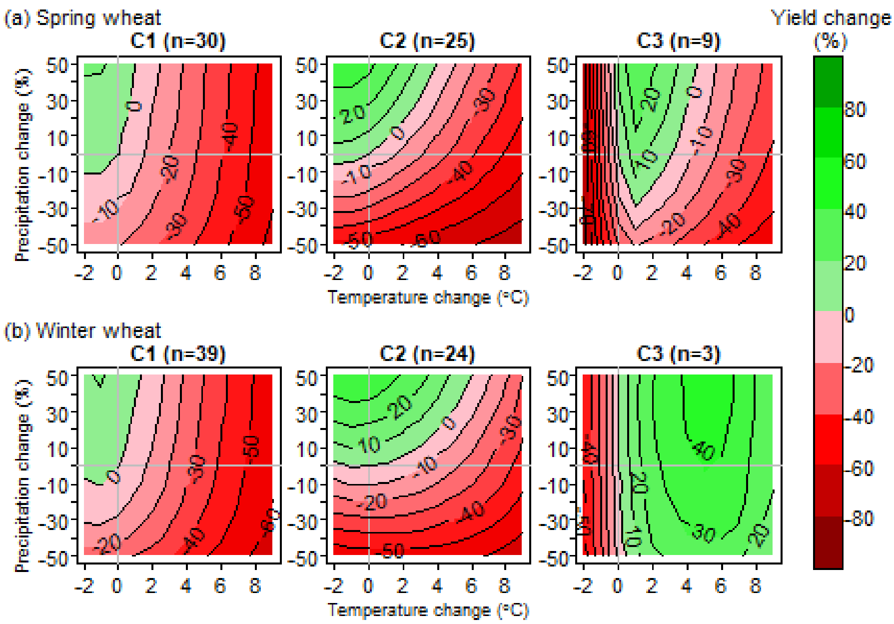

Figure 12 IRS patterns of thirty-year mean grain yield changes relative to the baseline (%) averaged across all members of clusters determined using the statistical diagnostic approach (SDA) for spring wheat (top row) and winter wheat (bottom row) for clusters 1– 3. The cluster identifier is indicated above each plot with the number (n) of cluster members in parentheses. For details, see text.

At the different sites there was some variation in the most frequent pattern as defined by EDA and SDA (see Figure 7 in paper **II**). At the German and Spanish sites, for both crops and for winter wheat in Finland, the most frequent patterns were of types C1 and C2, with the yield maximum at temperatures below or close to the baseline. Responses of the WOFOST model for spring wheat in Finland were similarly clustered into C1, whereas the most frequent pattern for spring wheat in Finland across the ensemble was C3 with a slightly higher temperature optimum for the yield.

The metric of distance (d) between IRS patterns applied in defining the clusters in the SDA was also used to provide a measure of similarity of the response patterns within a group of models defined by a common property. Thus, it offered the possibility to investigate whether the similarities/differences in the yield response patterns could be related to shared/diverging properties of the crop models. Figure 13 shows distance ratios for different model property types, ranked separately for groups defined by site, genealogy, i.e. belonging to the same family of models as interpreted here, and model calibration (upper panel) and by process description (lower panel). The measure indicates how similar the IRSs in each property group are compared to randomly selected groups of IRSs. For details see Paper **II,** subsection 2.4. Here, the figure is ordered for spring wheat rather than winter wheat (as in paper **II**) as this might offer insights that are closer to those of spring barley, not simulated by the model ensemble.

Influence of the site can be seen in the IRSs for Germany, for both crops, model behaviour being more similar than would be expected in random samples. In contrast, at the Finnish site the patterns for both crops were more dissimilar than if selected at random. At the Spanish site there was much more similarity among IRS patterns for winter wheat than for spring wheat. Models sharing common genealogy showed greater similarity in IRS patterns than those not characterised into any model family. With respect to calibration, the fewer the parameters involved in the calibration, the more similar the responses were in the defined groups with winter wheat. For spring wheat, the results on calibration were mixed.

With respect to the different process descriptions the interpretation is less clear. Note, that the number of members per group varies greatly with the range extending from 3 to 63. The results with respect to root distribution and treatment of water stress indicated that patterns were more similar between models having a simpler process description than with models having a more complex description. The relationship between model properties and the most frequent EDA class per property type found winter wheat to have a stronger precipitation response than spring wheat (not shown, see paper **II,** Figure 7), which was also seen from the classification of the ensemble median patterns (see Table 3).

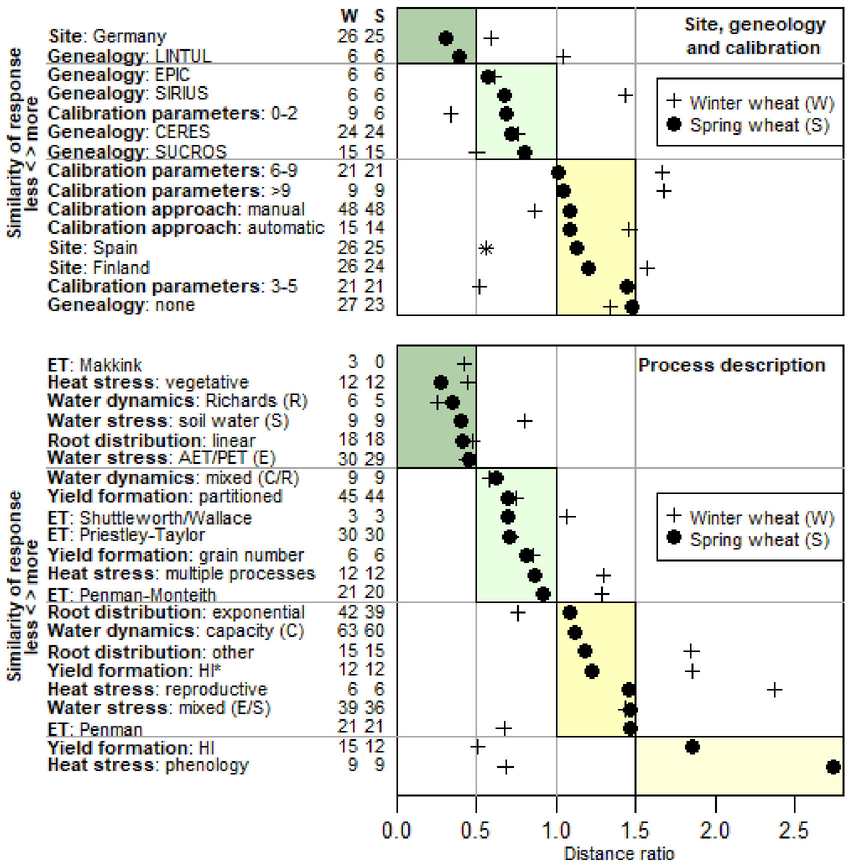

Figure 13 Ratios of the mean distance of the statistical diagnostic approach (SDA – see text), which compares the similarity of patterns between pairs of IRSs within a group of models sharing a specific property type (see Table 1 of Paper **II**), to the same measure computed for pairs randomly selected 100 times into groups of the same size from all IRSs of the model ensemble. Ratios are computed separately for winter (W) and spring (S) wheat and ordered from more similar to less similar groups for S (filled circles) for groups defined by site, genealogy, and calibration (top panel) and by process descriptions (bottom panel). Additionally, ratios are shown for the same groups for winter wheat (+). Group sizes are given under W and S. Lines and shaded rectangles delineate model groups with IRSs that are very similar (distance ratio ≤0.5, green shading), more similar than random (0.5–1.0, light green), more dissimilar than random (1.0–1.5, yellow) and very dissimilar (≥1.5, light yellow). HI=harvest index, HI*=modified HI).

3.4 APPROACHES TO ESTIMATING LIKELIHOOD OF YIELD SHORTFALL THROUGHOUT THE 21ST CENTURY

For computing likelihoods of 30-year mean yield shortfall, we adopted a threshold of 6000 kg ha $^{-1}$, which is the 30-year baseline mean yield simulated by WOFOST for spring barley cultivar Scarlett. It is used as an example of a hypothetical average yield level which farmers would not wish to see decline in the future. Likelihoods were computed using the IRS-based approach and applying a seasonal cycle for perturbing baseline temperature and precipitation throughout the year. By default, [CO_2] was assumed to evolve throughout the 21st century as projected by RCP8.5. For gaining confidence in the results received, the results of this chosen approach were compared against a more traditional approach running individual model simulations for each of the projected scenario climates.

In comparison to yields simulated using the scenario-based approach, the 30-year mean yields from the IRS-based approach, as interpolated to the same resampled GCM projections of temperature and precipitation change (Figure 14a), are in general slightly underestimated. However, a higher proportion of estimates fall below the threshold yield for the scenario-based than the IRS-based approach. This in turn translates into a slight under-estimation of the likelihood of yield shortfall (on average by 4% across the 21st century – see Figure 4 in Paper **III**) with the IRS-based approach even though the yields in general are at a slightly lower level. The colour coding of the points in Figure 14a shows again, rather consistently that the higher the temperature increase the lower the period mean yield.

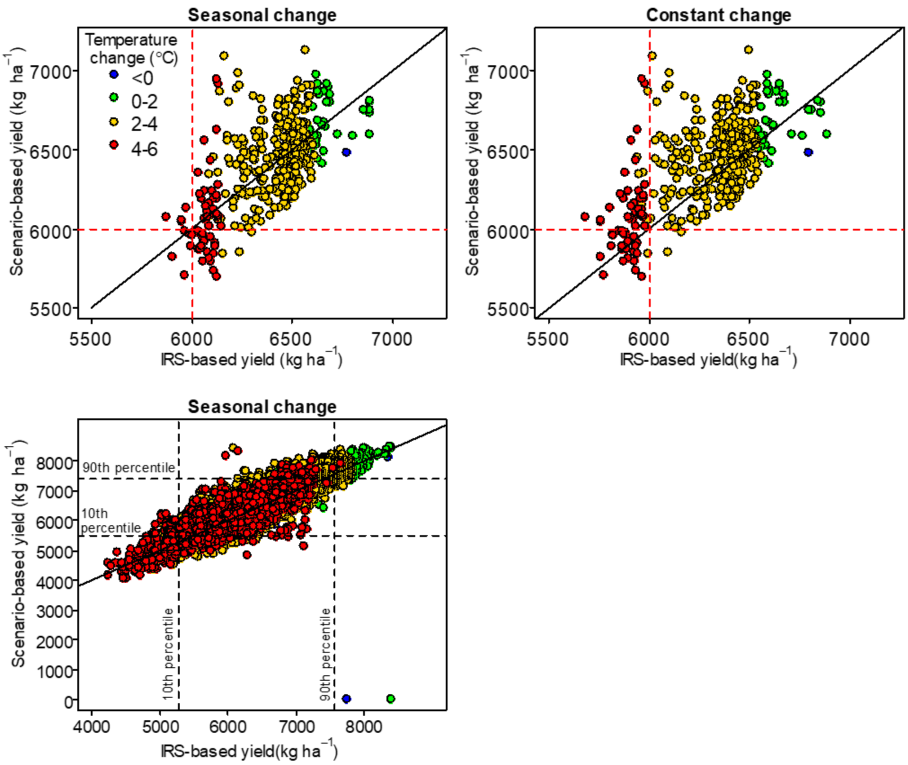

Figure 14 Comparisons of spring barley dry matter grain yields (kg ha $^{-1}$) at Jokioinen, Finland as simulated by WOFOST for 2041–2070 and 572 ppm [CO_2] (RCP8.5) using the scenario-based and IRS-based methods for combinations of resampled GCM period-mean temperature and precipitation projections (n = 378). Yields are plotted as 30-year mean yields simulated using (a) the seasonal change and (b) constant change methods and (c) as yearly yields using the seasonal change method (n=11340). In (a) and (b) the colours indicate the level of warming of each resampled GCM projection of temperature and precipitation change over Finland relative to baseline (dots) and the red dashed line is the 6000 kg ha $^{-1}$ yield threshold applied in calculating likelihoods of yield shortfall. The diagonal grey line in each plot is the 1:1 line of yields on the x- and y-axis.

Overall, for both approaches, with time-dependent [CO_2] increasing throughout the century the likelihood of yield shortfall for Scarlett slowly declines to zero from between 15 and 20% in the first future period centred on 2025, as increasing [CO_2] more than compensates for the negative effect on yield associated with warming. This contrasts with a rapid approach to 100% likelihood of yield shortfall when applying fixed [CO_2] throughout the century for the same warming (not shown – see Figure 4 in Paper **III**).

With respect to different methods of perturbing the baseline climate, applying a seasonal pattern as opposed to a constant change throughout the year results in slightly higher yields which shift the points more closely around the 1:1 line. Also, the scatter is slightly reduced (compare Figures 14a and b).

By plotting yield estimates similarly for the 30 individual years in each scenario climate (Figure 14c), it is found that while there is some scatter around the 1:1 line and the yields of the IRS-based approach are slightly underestimated, there is no systematic bias in the yields. The biggest differences in yields between the two approaches are seen for two projections where the scenario-based yield is zero, while for the IRS-based approach they are among the highest yields. This is due to the harvest cut-off in relation to the timing of the growing as defined by the dynamically defined sowing date (see sub-section 2.3). Here, sowing in the scenario-based simulation takes place over a month later causing the crop not to reach maturity before the cut-off.

3.5 CEREAL YIELD RESPONSES TO ADAPTATION OPTIONS UNDER A CHANGING CLIMATE

Adapting the sowing date of the crop according to temperature advances the time of sowing at Jokioinen on average (across 1981–2010) by approximately 10 days for every 2 degrees of warming. Under the baseline the period-mean sowing date is defined to be DOY 134 (mid-May), for a cooling of -2 °C it is DOY 146 (end of May) and for +8 °C DOY 97 (early April). The effect of adapting the sowing date shows as higher average yields with warming in comparison to applying baseline sowing dates, the effect enhanced with warming (Figure 15). Under cooling, both options end in crop failure (implemented through the harvest cut-off) during many years. With adapted sowing, having the sowing dates ~12 days later than with baseline sowing, this happens in 17 years out of 30 causing the median yield to be zero. However, with the large spread extending to 7812 kg ha^{-1}, during the years when the crop reaches maturity the yields are high enough to bring the mean yield up to nearly 3000 kg ha $^{-1}$. With baseline sowing the relationship between the mean and median is the opposite, with 10 years having zero yield and the remaining 20 years having yields between 5715 and 7798 kg ha^{-1}. Consequently, the median yield is very high, while the zero yields bring the mean yield down.

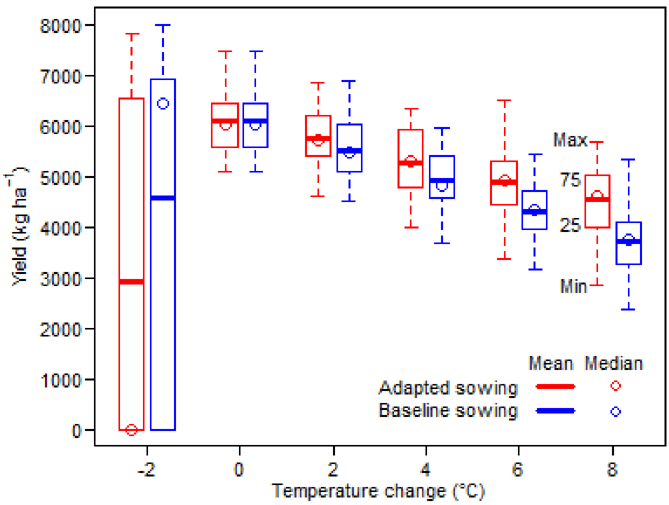

Figure 15 Modelled period-mean (1981–2010) yields of spring barley at Jokioinen for different temperature perturbations under baseline precipitation and assuming [CO_2] of 360 ppm using alternative ways of defining the sowing date. Red is for adapted sowing and blue for baseline sowing. Whiskers indicate the minimum and the maximum yields; boxes define the inter-quartile range (25–75 percentile). Mean and median yields at each temperature increment are indicated with a thick vertical line and an open circle, respectively.

Changing the cultivar to a slower developing variety was tested as another adaptation option. The temperature requirement of the growing period from anthesis to maturity (TSUM2) was found to benefit the yield levels more than that from emergence to anthesis (TSUM1; Table 4). The higher the requirement is, the slower the development and the higher the yield. Already during the baseline, currently used cultivars (Scarlett and Annabell) are outperformed with respect to the yield level by all cultivars, apart from the two slowest developing ones (Cultivars 9 and 10) that have the lowest mean yields due to occasional crop failures (Figure 6 in paper **III**). While Cultivar 7 also experiences occasional failures under the baseline, they are less frequent than with Cultivars 9 and 10, leading to a higher period mean yield (Table 4). When examining the yields across the century, the cultivars with the highest temperature requirement for TSUM2 (Cultivars 7-10) produced the highest yields from the first future 30-year time period onwards, centred on 2025 (Figure 6 in paper **III**).

Table 4. *Statistics of crop growth and likelihood of yield shortfall for three spring barley cultivars under baseline [CO_2] and climate (top section) and for projected [CO_2] and median climate change under RCP4.5 (middle section) and RCP8.5 (bottom section) for 2085. Statistics for RCP4.5 and RCP8.5 assume adapted sowing. Sow = sowing; doy = day of the year; Anth = anthesis; Mat = maturity; Failures = number of failed seasons during the 30-year period that each section is averaging; 1981-2010 for the baseline and 2071-2099 for 2085. In the coloured columns, orange highlighting indicates the highest value per column within each of the three sections separated by a dark grey divider.*

	Cultivar	Sow (doy)	Sow to Anth (days)	Anth to Mat (days)	Yield (kg ha^{-1})	Failures (n)	Likelihood of yield shortfall (%)
Baseline 1995 360 ppm Δt: +0°C δp: +0%	Scarlett	134	59	37	6082	0	-
	Cultivar 2	134	65	38	6087	0	-
	Cultivar 7	134	59	54	6659	3	-
RCP4.5 2085 532ppm Δt: +3°C δp: +10%	Scarlett	120	56	32	6239	0	19
	Cultivar 2	120	61	32	6092	0	34
	Cultivar 7	120	56	44	7444	0	1
RCP8.5 2085 802 ppm Δt: +5°C δp: +15%	Scarlett	112	54	31	6802	0	0
	Cultivar 2	112	60	29	6865	0	0
	Cultivar 7	112	54	42	8156	0	0

When analysing the likelihood of crop yield shortfall with respect to combinations of the adaptation options, at the end of the century all cultivars have zero likelihood of falling short of the defined threshold under RCP8.5 (Table 4; Figure 5 in paper **III**). Table 4 illustrates the relative importance of projected [CO_2] and median climate change at the end of the century, under RCP4.5 and RCP8.5, when adapted sowing is assumed. With RCP4.5 the yields are lower and likelihood of yield shortfall higher than for RCP8.5. By looking at the projected changes in temperature and [CO_2] for the two RCPs, it can be seen that the change in temperature compared to that of [CO_2] is relatively greater for RCP4.5 than for RCP8.5, which may explain the difference in the outcomes between the two. Across the century, the likelihood of yield shortfall for RCP8.5 is zero or very small with all cultivars, apart from the cultivars with the lowest temperature requirement for TSUM2 (600 °C d). These exhibit increased likelihoods of yield shortfall during the first half of the century, which thereafter diminish towards zero. Under baseline sowing the results are similar though the likelihoods of yield shortfall are higher for the faster developing cultivars (see Figure 6 in paper **III**).

4 DISCUSSION

The first two sub-sections of the discussion reflect on issues about model calibration, evaluation and uncertainties associated with the study, while the following sub-sections address the specific objectives of the research more directly. Finally, possible avenues for future work building on the results of the study are discussed.

4.1 MODEL CALIBRATION AND EVALUATION

In a multi-model ensemble study one of the most important premises for a successful study is ensuring the consistency and comparability of results produced independently by different modelling groups running models of varying complexity and requirements for input data. A carefully developed modelling protocol for the study and thorough checking of consistency of required choices made in running the model and input data used is essential. However, with a large group of models of differing complexity it is challenging if not impossible to ensure absolute comparability of results.

This is a particular concern with respect to model calibration. It is known that the choice of calibration approach has a significant effect on model results and thus contributes to the uncertainty in model results. The varied complexity in mathematical structure of models, with multiple inputs and outputs, difficulties in coupling models with existing calibration software and lack of existing guidelines on approaches to adopt has led to varying practices applied in calibrating different crop models (Seidel et al., 2018; Wallach et al., 2014). Further, there is a "user effect" in model calibration, due to subjective choices made. Differences can occur even when using the same model, sometimes resulting in uncertainties in impact estimation of even greater relevance than the way that knowledge is formalised in the model (Confalonieri et al., 2016). Another challenge is that of overfitting, where the use of too many variables in calibrating a model may reduce its predictive power due to the likely errors in the specifications of the model (Whittaker et al., 2010). Moreover, different combinations of parameter values (and multiple model structures) may lead to the same outcome and thus calibration may not result in unique parameter values – the concept of equifinality (Beven and Freer, 2001). Suggested improvements associated with calibration uncertainty include more and better data, easier to use calibration software and guidelines as to the choices made in relation to model calibration, among others (Seidel et al., 2018).

In papers **I** and **II** no specification was given as to how calibration should be performed, beyond provision of calibration data. As a result, the chosen approaches differed from no calibration at all (i.e. using an existing out-of-the-box calibration) to applying automatic methods where values of multiple parameters were adjusted. These choices in turn affected the results, though identifying and quantifying the effects is not so straightforward. Our results suggested greater similarities in response patterns with less complex calibration approaches for winter wheat but showed no clear tendencies with spring wheat in this respect. However, visual inspection of period mean IRSs produced by individual modelling groups per site and crop (Supplement 2 in paper **II**) shows differences in patterns produced by different modelling groups using the same model and study set-up (compare IRSs for ID 7 with 8 and ID 25 with 26). Thus, the difference in yield response can be attributed to differences in calibration and modeller preferences in applying the model.

Various considerations may affect model evaluation when comparing observed and simulated yields. For instance, lack of observed data for model evaluation at the required spatial resolution (e.g. sub-regional or regional level) may necessitate the use of data representing much larger areas, such as national statistics provided by FAO (Food and Agriculture Organization of the United Nations, 2017) as in papers **I** and **II**. Consequently, discrepancies between observed and simulated yields may largely reflect differences in climate or soils of the respective sets of data and not necessarily the capability of the model to simulate the relationships between input and output data such as those of weather and yield. The data available for model calibration provided site observations for only a few years. In paper **III,** the use of the official variety data, available for barley from the same location as used for performing the simulations, was hindered by the large spread in the yearly observations with relatively low numbers of observations per year. In summary, all three sets of observed data had their limitations, differing in resolution, time span and number of observations and hence complicating the evaluation of model results. However, the overall comparison highlighted issues such as the yield gap often associated with comparisons of observed and simulated yields, where simulated yields exceed the observed. While suggestions have been made on standard approaches with respect to the content and quality of data sets for purposes of crop modelling (e.g. White et al., 2013) and harmonised data sets have been collected and made available (Porter et al., 2014a), the need for more observations better suited to serve modelling on a local level still remains.

The difficulty of separating the influence of technological development and improved management practices from climate on yield may further complicate model evaluation. Under the assumption that new technologies, such as improvements in genetics and management associated with plant protection and application of mineral fertilizers, have affected crop yields more than any other factor since the green revolution up to the present (Crespo-Herrera et al., 2018; Evenson and Gollin, 2003; Voss-Fels et al., 2019), observed yields are often de-trended in an attempt to distinguish the variations in yields due to yearly weather conditions (Easterling et al., 1996). However, as a result possible trends in climate may also be removed along with the technological trends. On the other hand, for model evaluation under the baseline climate the focus is primarily on examining the capability of the models to respond to yearly weather variations.

In light of the baseline period of the study being 1981–2010, it is interesting to reflect on any changes in climate that have occurred since that time. At Jokioinen there have been increases in temperature between 1981–2010 and 1991–2019 of on average 0.3 °C during spring (March-April-May) and summer (June-July-August). The biggest increase in an individual month is that of 1.2 °C in December. The yearly pattern of change across seasons resembles that applied in the seasonal change approach for perturbing future temperature, with winter warming more than summer. Interestingly, for an increment of 0.5 °C of annual mean change, which equals the observed period mean annual change, the warming for winter is underestimated in the seasonal weighting pattern compared to what is observed (see Paper **III**, Figure S1). However, for spring barley, the crop used in the study, this has little relevance as winter is well outside its growing period. With respect to observed change in precipitation, there is a decrease in the autumn months (September-October-November, mean -4.4%) and an increase in winter (December-January-February, mean 4.2%) in comparison to the earlier period. During spring and summer, the direction of change between the two periods alternates from month to month between -7.7% and 8.4%. The period mean annual change relative to the study baseline (1981–2010) is -1.6%.

For analysing the effect of the observed changes in climate to yield, simulated spring barley yields were interpolated to a re-sampled GCM projection with very similar projected changes (0.5 °C for temperature and -1.1% for precipitation) as observed between 1981–2010 and 1991–2010. It was found that the yield declined very slightly, if the increase in [CO_2] was not accounted for but taking the increase from the 1995 [CO_2] of 360 ppm to the [CO_2] of 2005 (380 ppm – mid-years of both time periods) into account, the yield decline was more than compensated.

4.2 UNCERTAINTIES IN CROP YIELD ESTIMATES

In the studies presented here the main approaches to addressing uncertainty in crop model estimates involved the use of a crop model ensemble in papers **I** and **II** and projections of climate change interpreted probabilistically in paper **III**. In another multi-model study, comprising seven crop models for barley at the same Finnish and Spanish locations and looking at uncertainties in yield estimates for the 2050s, crop model structure was found to contribute more to the total variance of the ensemble output than downscaled climate projections and model parameters (Tao et al., 2018). The ensemble of models applied in papers **I** and **II** was larger, incorporating crop models of various complexities and scope for application. Their common feature is that they are all documented, process-based models, capable of simulating effects of climate change on crop yield. The attempt made to analyse this structural uncertainty stemming from differences in model structure involved analysis of the IQR in yield responses between models. Results suggested greater spread in simulated responses for certain extremes with respect to both temperature and precipitation that were site- and variety-specific. This was interpreted as reflecting the differences in how models simulate various challenging conditions, such as water-deficit, extreme heat effects and over-wintering in winter wheat, and how they translate these to differences in yield responses.

Through the quantitative representation of uncertainties, probabilistic climate projections have been found to open useful avenues for research when used in connection with impact models. Over the years there have been various attempts to quantify uncertainties in future climate projections in a probabilistic way through use of multi-model ensembles that sample initial, parameter as well as structural uncertainties in climate model design (Tebaldi and Knutti, 2007). Among the first to propose constructing a probabilistic view of climate change projections on the basis of multi-model estimates were Räisänen and Palmer (2001). The method used in paper **III** is based on a relatively simple resampling method, relying on the CMIP5 ensemble of climate model simulations. It results in an increased sample size of projections of future temperature and precipitation change, comparable in magnitude over Finland to those obtained using more complicated approaches combining perturbed physics experiments with multi-model ensembles (Harris et al., 2010). It should be noted that, while calling these projections probabilistic, capturing the “true”, objective, probabilities of the different scenarios and projected changes is unachievable. As such, the use of the term may be contested but the projections used here aim for an approximation of the associated probabilities of what is known, through use of subjective expert judgement, to allow for any impact risk assessment to be conducted (Gay and Estrada, 2010). Furthermore, such probabilistic projections are always conditional on a given forcing scenario, with specific underlying assumptions attached. Thus, projections associated with different RCPs cannot be combined to include this aspect of uncertainty.

A question remains about the effect of possible changes in the inter-annual variability in climate which is not built into the estimates applied in paper **III** that are based on mean annual changes alone. In the estimates reported here, inter-annual variability is conveyed only through the variability portrayed in the baseline climate.

Among the limitations associated with such probabilistic representations of climate change is the nature of the models making up the ensembles, being essentially “ensembles of opportunity”, similarly as the ensemble of crop models applied in papers **I** and **II.** What this implies is that the ensemble of models is constructed from all models for which there are comparable simulation results that are available for scrutiny. As a result, the size and composition of any ensemble is determined based on aspects related to interest, funding and resources rather than on an attempt to capture the uncertainty across models based on either systematic or random sampling (Tebaldi and Knutti, 2007). Suggestions have been made for improving the rigour in ensemble modelling, for example, through defined criteria for acceptance of models into the ensemble and studying the effect of number of models in an ensemble (e.g. Rodríguez et al., 2019; Wallach et al., 2016).

With respect to the uncertainties associated with the RCPs, one aspect is that the four RCPs were originally selected to embrace a representative range of uncertainties in the future development of factors affecting the radiative forcing of the climate (such as atmospheric composition and land use change) reported in the literature. No probabilities can be attached to these idealised pathways. Projections of climate change that are based on the RCPs are themselves uncertain, due to the incomplete understanding of the climate system represented in models used to simulate the climate response to forcing and internal climate variability (Collins et al., 2013). In paper **III**, although simulations were performed for RCP4.5 and RCP8.5, the low end of the climate response to forcing is not included, typically represented by RCP2.6 (van Vuuren et al., 2011). While not spanning a wider range of uncertainty in radiative forcing, the simulations for the two RCPs did reveal interesting results on the relative importance of the increasing [CO_2] in combination with increasing temperature. Specifically, lower increase rates of [CO_2] fail to compensate fully for the yield losses due to concurrent warming under RCP4.5 for the fastest developing cultivars. In contrast, under RCP8.5, likelihoods of yield shortfall for barley fall to a lower level than with RCP4.5 from mid-century onwards and eventually approach zero. However, this result needs to be interpreted with care, particularly with respect to areas relying on irrigation, as the challenges for adaptation, through more extreme climate events and bigger changes in growing conditions, are still likely to be more severe for the more extreme scenario of RCP8.5 (Levis et al., 2018).

In addition to the uncertainties discussed above and the uncertainty associated with model calibration, the highly stochastic nature of processes associated with agroecosystems adds uncertainty to input data with variation arising from issues such as farmer behaviour, machinery performance, spatial variability of individual fields and effects of pest, weeds, diseases and unfavourable harvest conditions. These effects are typically not captured in individual models and may thus lead to differences between observed and simulated yields. Uncertainties are also associated with soil profile descriptions due to within field variability and measurement error (White et al., 2011). Here, two somewhat extreme alternative soil types were used to inform about the sensitivities and uncertainties associated with the choice of soil type in the simulations with WOFOST. However, as the soil description in this particular model is a relatively simple one, the results were discussed only on a rather superficial level.

4.3 SENSITIVITY OF CEREAL YIELDS TO CLIMATE CHANGE

This study has analysed the sensitivity of winter and spring wheat and spring barley varieties to changes in temperature and precipitation (objective 1). In addition to providing insights on the response of yield and behaviour of models across a range of changes in climate (objective 2), the examination of sensitivity also provides information of the vulnerability of cereal yields under long-term climate change, a component of risk assessment. Vulnerability is interpreted here in physical terms as the inherent sensitivity of the crop to given changes at a given site (defining the hazard).

The sensitivities identified for spring varieties of cereals at each site along the latitudinal transect (based on the EDA), reflect the baseline climate conditions. The temperature dominated response in Finland results from the constraints posed mainly by temperature on the growing season, while precipitation rarely limits yield. In Spain, water is already a major limiting factor. This is reflected also on the IRS as a strong response to precipitation changes, making the crop vulnerable to future decreases in precipitation. In Germany, on average the conditions are more favourable. On the IRS this shows as a weak response with respect to both drivers.

On the differences between winter and spring wheat, in a study by Peltonen-Sainio et al. (2011) based on long-term field experiments with existing varieties, winter wheat in Finland was found potentially to benefit from increased temperatures. In this study, a yield decline with warming was found with both wheat varieties, though the decline was weaker for winter wheat, especially in Finland and Germany where spring wheat is sown in the spring and winter wheat in the previous autumn. Following over-wintering, winter wheat matures on average earlier than spring wheat, exposing it to more optimal growing conditions than those often found at the end of the season with higher temperature and occurrences of drought. In Germany the growing period of winter wheat is also longer than with spring wheat, allowing more time for yield formation.

The yield decline with both cooling and warming observed with both wheat and barley reflects the fact that the cultivars are bred to perform best under ambient climate. Cooling results in crop failure if the crop cannot mature in time. Warming accelerates plant phenological development, leaving less time for allocation of dry matter to the grain which translates to lower yields (Kontturi, 1979; Peltonen-Sainio et al., 2011; Rötter et al., 2011b). With respect to changes in precipitation the main effect is through the limitation soil moisture deficit may cause on growth. With increased precipitation the deficit is alleviated and with warming, through increased evapotranspiration, and drying the effect is more pronounced. However, with increasing [CO_2], through its beneficial effects on crop growth and water use efficiency, the yield losses are compensated to some extent, as also illustrated with the results for the two periods of observed weather (see sub-section 4.1). The higher the

[CO_2], the less the decline in yield is with warming. In climatic regimes such as in Finland, where precipitation is less of a limiting factor, there are only limited gains to be realised with increases in precipitation, especially with favourable soils such as clay loam where water deficit is rarely experienced. The specific processes associated with climate impacts on crops are discussed in detail in papers **I** (wheat) and **III** (barley). The lesser sensitivity of barley than wheat to warming reflects the difference between single model output and the mean response of a multi-model ensemble, as well as aspects associated with calibration and model set-up, rather than any crop specific responses. When running WOFOST for both crops, using the same modelling set-up, the responses were found to be very similar.

The optimal temperature for local cultivars was found to be close to the baseline under Finnish conditions, while at the German and Spanish sites yields benefited from some cooling. This suggests that the adoption of cultivars with higher temperature requirements for development might already be beneficial at these sites, and even more so under projected warming. The yield declines found with warming are largely consistent with both observed yield trends worldwide (Lobell and Field, 2007) as well as with previous multi-model studies for constant [CO_2] (Asseng et al., 2014a; Asseng et al., 2013). When CO_2 is accounted for, as with the simulations for spring barley in Finland, a yield increase of 12.5% is found for the current cultivar Scarlett at the end of the century for RCP8.5 where the [CO_2] is 802 ppm and the multi-model median temperature increase is 4.9 °C. Other studies have also found increases in cereal yields with evolving [CO_2] accounted for, though there is a lot of variation in the models, scenarios and target future periods applied, hindering exact comparison of results (e.g. Iglesias et al., 2012; Olesen et al., 2007; Trnka et al., 2004).

The results from the classification of yield responses (based on the SDA) and the spread of the IQR around the multi-model ensemble mean at different sites reveal that the simulations under conditions experienced in Germany result in the most similar response patterns. These findings could reflect the fact that on average there are less stresses to account for in the modelling with respect to heat and drought but also to cool conditions. While drought stress is projected to remain as the main driver of crop losses (Webber et al., 2018), extreme heat events, which exceed the maximum temperature limits of specific growth stages above which growth ceases, are already observed occasionally at all study sites, though they are more pronounced and frequent in the warmer and drier climates, such as that observed in Spain. The treatment of such extreme events is variable across models, many lacking representation of heat stress effects, for instance, or accounting for them is done inadequately, even though research shows the importance of treating such effects in crop models (e.g. Liu et al., 2016).

A measure of yield reliability (percentage of years when yield is above the 10^{th} percentile of the baseline yield) and CV were used for analysis of the inter-annual variability in wheat yields under changes in temperature and precipitation. Due to the method of perturbing baseline weather data for constructing an IRS, the results essentially reflect the yearly variation in the baseline weather and how the yields respond to changes in it. Generally, at all sites, reliability is increased where mean yields are highest. With a decline in mean yield, when the general yield level is reduced, yields in more years fall below the threshold and reliability is similarly reduced (see Paper **I,** Figure 7 and Figure S3).

The difference between the patterns of reliability and CV reflect the fact that CV accounts for the full yield distribution, while reliability as defined here focuses only on the lowest yields. While the level of period mean yield closely relates to both measures, the patterns themselves have less of a resemblance (see Figures 5-7 and Figure S3 in Paper **I**). The exact mechanisms of the underlying causes leading to the patterns of CV are less straightforward to disentangle and conducting such a study was not found to be feasible within the resources and scope of this study. However, it is reasonable to assume that the weak response of wheat to temperature changes in Germany reflects greater stability in yields from year to year across a wider range of temperatures as in Finland where CV increases rapidly to either side of the temperature space with the lowest CV (around +3°C). This is caused by occurrences of very low yields or crop failure during some years while some years still produce relatively high yields. The large contrast between CV patterns of spring and winter wheat in Spain, with rather opposite response to the two drivers, is assumed to relate to both spring and winter wheat being sown in the autumn and to the vernalisation requirement of winter wheat that is included in most, but not all of the models.

4.4 LIKELIHOODS OF YIELD IMPACTS AND EFFECT OF ADAPTATION OPTIONS

Choice of impact threshold was found to influence the received results on likelihoods of yield shortfall quite considerably. For example, setting a very low yield threshold would have resulted in zero likelihood of yield shortfall throughout the century. To assure the relevance of the received results for practical purposes, the choice of the impact threshold for estimating likelihoods needs to be carefully considered and ideally defined through consultation with stakeholders, inevitably involving selective judgement (see discussion by Lachaut and Tilmant, 2020). However, to allow examination of differences between, for example, alternative adaptation options, it may also be justifiable to choose the threshold in an exploratory context, like in paper **III**.

The measure of likelihood of yield shortfall was used in this study as a means of attaching probabilities to yield outcomes in the context of risk assessment, specifically with respect to the efficiency of adaptation options in affecting the likelihood of shortfall (objectives 3 and 4). The focus of the options was on increasing the resilience of crop production at the local level through changes in sowing time and adoption of cultivars capable of making better use of the growing season within the constraints set by the local climate as projected for the future. In the risk context this can be framed as altering the exposure of the crop to the hazards posed by the climate, although the emphasis is more on exploiting the benefits of the changes in the growing season as opposed to only avoiding the negative consequences.

The study of paper **III** on spring barley yields in Finland found that the combination of adapted sowing and a cultivar with higher temperature requirements for crop development resulted in the highest yield gains and lowest likelihoods of mean yields falling short of a predefined threshold. With warming, adapted sowing results in earlier sowing dates, which lengthens the potential growing season, while higher temperature requirements allow the crop to take advantage of the longer season with slower crop development, thus leading to higher yield potential (see also Liu et al., 2018). The temperature sum requirement during the reproductive phase, between anthesis and maturity (TSUM2), was found to be the determining factor for achieving the greatest benefits of projected warming. Higher values for TSUM2 correspond with higher yields and lower likelihoods of yield shortfall. Conversely, a lower temperature requirement for TSUM1 (emergence to anthesis) was found to be beneficial, though the effect of this was much weaker than that of TSUM2. Similarly, Tao et al. (2017) found for the same crop and location that a shorter pre-anthesis phase and a longer reproductive phase were among the most effective traits in high-performing barley cultivars under future climate. The longest developing cultivars were found to offer the highest and most stable yields from 2035 onwards. Up to then the slow development might still lead to more frequent crop failures than for current cultivars, leading to lower yield reliability.

A possible explanation for a lower TSUM1 requirement being advantageous is that the longer the pre-anthesis phase is, the more the crop is exposed to the period with the highest temperatures and longest days with high radiation potential. Such conditions lead to hastened development rate and intense growth, associated with yield penalties during pre-anthesis (Peltonen-Sainio and Rajala, 2007). A higher TSUM2, on the other hand, extends the grain filling period when assimilation is directly translated into grain yield, thus leading to higher yield potential (Olesen et al., 2012).

While paper **III** focused on the effects of adaptation options for spring barley at one site in Finland, the effective adaptation options are very specific to local climate conditions. For example, at the Spanish site Lleida it has been found that while early sowing is similarly beneficial as in Finland, it is a shorter growth cycle that offers most promise through avoidance of the negative effects of high temperature at grain filling. Extreme high temperatures at the end of the growing season with longer duration cultivars could also impose an absolute constraint on wheat viability. Furthermore, in Lleida, switching to spring wheat varieties from the currently preferred winter wheat was found beneficial as the vernalisation requirement in winter wheat is delayed or may not be met under warming temperatures (Ruiz-Ramos et al., 2018).

Breeding of new crop cultivars addresses several other genetic traits such as improved stress tolerance, resistance to pests and diseases and aspects associated with photosynthesis and grain formation (Tao et al., 2017). However, modelling offers a powerful tool for assisting in the breeding process through the possibility of testing a wide range of options and identifying the most promising traits for actual breeding experiments as well quantifying possible future threats to crop production (e.g. Semenov and Halford, 2009). It should be noted that in actual breeding efforts various physiological links and constraints of crops need to be considered, for example relating to the specific durations of different growth phases. Thus, in reality lengths of individual growth phases cannot freely be varied independently to the extent assumed in this analysis of the sensitivities associated with the TSUMs (FAO, 2002).

Both options for adaptation studied here can be regarded as autonomous, as farmers have always adjusted their practices according to expectations of the coming weather. The simulations without any adaptations assumed represent a less plausible "dumb farmer" approach, where practices remain unchanged regardless of changes to climate (Easterling et al., 1992). Equipped with information about the impacts of future climate and having available daily local weather forecasts extending nowadays up to 15 days ahead with improved accuracy due to advances in technology and prediction methods, farmers are better equipped to adapt and plan their activities according to yearly conditions. According to Kaukoranta and Hakala (2008), farmers were found to have adequately adjusted their field activities along with changes in spring temperature with sowing of spring cereals having advanced 2–2.8 days per decade largely due to warmer springs. In our simulations the sowing date was estimated to be 10 days earlier for every 2 °C of warming equating to 2.5 days for 0.5 °C which was approximately the change in temperature between 1981–2010 and 1991–2019. In addition, to the options tested in this study, irrigation is often included in crop modelling studies of the efficacy of adaptation and found to offer promise in overcoming some of the detrimental effects of climate change, particularly in drier environments already challenged by water deficit (e.g. Ruiz-Ramos et al., 2018).

4.5 MODELLING SET-UP

When performing a modelling study for analysing impacts of climate change on a chosen impact variable, such as crop yield, the model and study set up can have a crucial effect on the results obtained (objective 5). In a model ensemble study with multiple modelling groups, constraining the number of options to be tested (alternative crops/soils/methods of defining the sowing date etc.) limits the number of simulations to be performed, thus likely increasing the participation rate of modelling groups and decreasing the likelihood of errors in applying consistent settings for the model runs across the ensemble. For an individual running a single model there is more room to explore different options for setting up the model and the study, thus allowing for more of the uncertainties and sensitivities associated with the study to be addressed.

4.5.1 CHOICE OF MODEL

The IRS-based approach was applied in both study set-ups to estimate crop yields and in paper **III** likelihoods of yield shortfall. Among the main differences in the approaches was the use of a single model, WOFOST, in paper **III** as opposed to using an ensemble of crop models in papers **I** and **II.**

Based on the analysis of the results of the model ensemble, WOFOST produces yields that are towards the upper range of grain yield estimates. With a favourable soil, such as the clay loam used, it is relatively insensitive to precipitation changes, thus reducing the inter-annual variation due to dry years. For the same model with coarse sand, having weaker water holding capacity, the yearly variation is increased, especially during drier growing seasons. This insensitivity with clay loam may be due to the relatively simple description of soil processes in the model and/or how the model was calibrated with respect to the soil parameters. On the other hand, the model being designed for simulating crop growth with detailed descriptions for phenological development, leaf related processes and CO_2-assimilation (de Wit et al., 2019) makes it a valid candidate for analysing impacts of changing climate and increasing [CO_2] on crop productivity and yield in common with other similar models of the class (e.g. Rötter et al., 2011a; White et al., 2011). As an improvement in future work, WOFOST could be coupled to a more advanced soil water balance model than the tipping bucket approach originally applied with the model, so as better to keep track of the moisture content in the soil and thus improve the treatment of the precipitation response (de Wit et al., 2019). One example of this is the Soil-Water-Atmosphere-Plant (SWAP) system, which uses WOFOST to simulate annual growth of crops and grasslands (Kroes et al., 2008; Kroes et al., 2000).

Considering the variation in the yield responses to changes in temperature and precipitation across individual models, the use of a model ensemble, as in papers **I** and **II,** can offer both a "consensus" view of the average yield response and information on the uncertainty in the results across a range of models. The use of a model ensemble has been promoted to increase the robustness of modelled yield estimates rather than relying on individual models. This is because in an ensemble errors in process descriptions tend to cancel each other out and the coverage of knowledge and different processes is wider than in any individual model (Angulo et al., 2013; Asseng et al., 2014a; Martre et al., 2015; Wallach et al., 2018). Further, many additional benefits arise from the close collaboration between modelling groups that can aid the sharing of information and development of ideas, to name a few (Wallach et al., 2016).

It should be noted that there are still processes, such as those associated with pests, diseases and various extreme weather events, that are not captured in any of the models included in the ensemble used in this study. Current yield losses due to pests, as estimated in a simulation study over 2001–2003, already account for nearly half of the attainable production across Europe, with crop protection found to be capable of recovering approximately two-thirds of these losses in north-western Europe, about a half in north-eastern Europe and about one-third in south-western Europe (Oerke, 2006). Single and compound extreme weathers, causing heat shocks, excessive rainfall effects, such as flooding, stem bending and erosion have also been found to impact yields negatively (e.g. Beillouin et al., 2020). However, due to complexities associated in simulating the effects of pests, diseases and weeds and such extreme weather events, these are not treated explicitly in this study.

4.5.2 UTILITY OF THE IRS-BASED APPROACH

In paper **III**, the likelihood of crop yield shortfall was estimated by utilising resampled projections of temperature and precipitation change. Results obtained using the IRS-based approach, requiring a single WOFOST-based, 30-year averaged IRS, for each future time period, were evaluated against the more conventional method of deriving likelihoods. This involved conducting WOFOST simulations for each resampled climate projection in each future time period (scenario-based approach). Here, the main difference between the two approaches is how the seasonal pattern of change in temperature and precipitation is taken into account. In the scenario-based approach, the individual seasonal patterns of temperature and precipitation changes as portrayed by each resampled GCM-based projection are accounted for. The IRS-based crop model simulations assume a generalised seasonal pattern, defined from the GCM-based projections, for perturbing the two climate variables. Compared to assuming fixed annual changes to perturb climate, the addition of seasonality produced higher yields, bringing the results closer to those estimated by the scenario-based approach. This is because summer temperatures for a given annual change during the growing season are lower than when applying a fixed annual change throughout the year, thus prolonging the critical growth phases and allowing for higher yields to be developed. On the other hand, when precipitation is limiting, the yields are lower with the seasonal changes as an annual average increase in precipitation converts to a summer decrease. However, as water deficit rarely limits yield in the simulations of this study for Finland, this has hardly any effect in the period mean results. The importance of accounting for seasonal differences in climate changes has also been shown in earlier studies (Børgesen and Olesen, 2011; Fronzek et al., 2010; Wetterhall et al., 2011).

Overall, only minor biases in yield level and impact likelihood estimations are introduced when using the IRS-based approach (with seasonal changes), making it a credible, simplified alternative to the computationally more demanding scenario-based approach, which would need to be repeated in full every time new scenarios become available. It is indeed this “scenario-neutral” nature of the IRS-based approach (Prudhomme et al., 2010) that is among its key benefits. As long as the projected changes of possible new scenarios are within the ranges used for constructing the IRS and all assumptions for running the impact model remain the same, yields estimated by the IRS-based approach can be attached to any future projection of the same two variables as used for constructing the IRS. Typically, regionalised scenarios of climate change are directly applied to simulate the effects of these changes on the system being modelled (e.g.Hoegh-Guldberg et al., 2018). Similarly, as in the scenario-based approach applied here, each impact estimate is then conditional on the scenario being used. Thus, to allow exploration of a range of alternative future conditions, particularly in an attempt to capture some of the uncertainties associated with the climate projections, a very large set of scenarios is needed – very likely larger than the number of impact model simulations needed for constructing an IRS (which depend on the number of perturbations to the baseline included).

In the context of evaluating adaptation options, the difference between the IRS-based and the scenario-based approach is in the sequencing of the stages of analysis. The scenario-based approach is often also referred to as the “top-down” or hazard-oriented approach (see sub-section 1.2), which starts with the climate scenarios to assess how the climate will change. Once impacts are assessed with respect to given scenarios, adaptation can be planned and evaluated. Uncertainties propagate at each step of analysis and are highly scenario dependent. A more systematic method for identifying appropriate adaptation measures approaches the issue from “bottom-up” or from a vulnerability/exposure-oriented perspective. The idea here is to address first vulnerabilities and sensitivities of the system or entity at hand in a rigorous and systematic manner for producing a more robust assessment of key uncertainties. This may then serve better the purposes of identifying key thresholds of impact that if exceeded might require adaptive responses. Climate scenarios are only brought in later to evaluate the likelihood of such exceedances occurring to help select the appropriate adaptation responses (Falloon et al., 2014; Lal et al., 2012; Lavell et al., 2012).

Additional benefits of the IRS-based approach relate to its presentational aspects. It facilitates the visual interpretation of yield response and crop model behaviour across a large range of plausible changes to the two driving variables. In doing so, possible discontinuities in the model response and issues in model behaviour are revealed that can assist in model testing. In the context of the multi-model ensemble simulations for papers **I** and **II,** modelling groups were given the opportunity to rerun their simulations after examination of initial IRS plots, sometimes revealing needs for refinements in the model set-up. The possibility to present visually how the uncertainty in the climate projections changes and shifts on the IRS through time and translates to the likelihood of yield shortfall when related to a threshold yield, provides better understanding and improved transparency of the concepts behind the results than presenting likelihood values alone. This possibility to relate impacts to probabilistic projections of climate change also facilitates consideration of adaptation – both its urgency with respect to the changing climate and, if they can be simulated, the effectiveness of different adaptation measures in avoiding the exceedance of critical thresholds. Use of common variables such as temperature and precipitation, allows easier comparison of results across models, regions, and impact sectors (Fronzek et al., 2018).

While the idea of applying IRSs in conjunction with probabilistic projections of climate change was introduced already two decades ago (Jones, 2000), it has only gained popularity more recently. Aspects of the approach have since been explored, for example in the context of wheat yields (Børgesen and Olesen, 2011; Ferrise et al., 2011; Luo et al., 2007), permafrost habitats (Fronzek et al., 2011; Fronzek et al., 2010), forest fires (Mäkelä et al., 2014) and water resources (Holmberg et al., 2014; Kay et al., 2014; Lachaut and Tilmant, 2020; Weiß, 2011; Wetterhall et al., 2011). The effectiveness of adaptation options for wheat has also been examined through IRS analysis (Rodríguez et al., 2019; Ruiz-Ramos et al., 2018), though not in combination with probabilistic climate projections to address aspects of impact likelihoods. To our knowledge, the study reported in paper **III** is the first to combine both the probabilistic assessment of impacts and assessment of adaptation options as a way of addressing climate related risks to crop yields.

In this study the decision was made to plot the IRSs as two-dimensional plots depicting the two drivers on the X- and Y-axis and the impact variable (yield) on the surface of the plot. Examples are also available on plotting the impact variable on a third Z-axis, showing the response as a three-dimensional graph (Allen Jr., 2019; Van Minnen et al., 2000; Weiß and Alcamo, 2011). Attempts have also been made at showing a third driver of the impact response on the Z-axis (Luo et al., 2005; Luo et al., 2007). While a three-dimensional plot could technically allow for simultaneous depiction of responses to an additional variable e.g. to changes in temperature, precipitation and [CO_2], interpretation is greatly complicated. Problems associated with interpretation of three-dimensional graphics relate partly to the relatively inferior ability of the human eye to extract and interpret information related to depth. Thus, visualisation of abstract numerical information rarely increases readability presented in three dimensions (Koponen and Hildén, 2019), a view shared by the co-authors of the three papers included in this thesis, which is why only two-dimensional surfaces are presented.

The two-dimensional nature of the IRS-based approach can be considered a limitation with respect to simulating impacts that are clearly dependent on more than two impact variables. Assumptions about how to treat other relevant explanatory variables in accordance with the perturbed variables need to be made, which may lead to simplified representations of the real-life relationships. However, the approach is not tied to the use of temperature and precipitation as the two driving variables. In principle, any driving variables to which an impact model is sensitive (climate or non-climate) could be used to construct an IRS (for an example for a range of sectors across Europe see Fronzek et al. (2018)).

To summarise, Table 5 lists the key features of the IRS-based approach and the scenario-based approach against which it was evaluated.

Table 5. *Summary of the key features of the IRS-based approach versus a more conventional scenario-based approach with respect to five different characteristics of the outcomes.*

Characteristic	IRS-based approach	Scenario-based approach
Representing future climate	Simplified	Detailed
	Key variables; others approximated	All relevant variables
	Scenario-neutral	Scenario-specific (opportunistic)
Representing impact model behaviour	Rapid appreciation of sensitivity to key climate variables	Responses to climate variables are scenario-specific and difficult to untangle
	Future climate uncertainties represented by systematic sensitivity analysis	Future climate uncertainties represented by ensemble model projections
	Rapid diagnosis of possible errors	Complex responses may obscure errors
	Allows rapid determination of impacts for any simple scenario	Scenario-specific responses; any new scenario requires new simulation
	Results sensitive to model set-up (e.g. calibration, [CO_2], soil, seasonality)	Results sensitive to model set-up (e.g. calibration, [CO_2], soil, variability change)
Ensemble impact model behaviour	Allows rapid visual and systematic inter-comparison between models	Reasons for inter-model differences may require post-processing of initial results
	IRS can depict structural uncertainties using measures of inter-model variability	Inter-model structural uncertainties can be represented but less easy to interpret
Response classification	IRS allows consistent classification of response patterns	Any classification would require analysis and post-processing of results
Impact likelihoods	Estimated by combining an IRS with climate projections interpreted probabilistically	Estimated with simulations for every climate projection and interpreting results probabilistically
	Threshold level of impact required	Threshold level of impact required
	Time-evolution of impact risks can be estimated and visualized with future IRS-climate combinations	Time-evolution of impact risks can be estimated and visualized based on future simulations for all ensemble members.
	Future changes in climate variability excluded or difficult to represent	Future changes in climate variability can be included

4.5.3 IMPLICATIONS OF REFINEMENTS IN THE MODELLING SET-UP

Additional differences in the two modelling set-ups relate to treatment of other critical concepts for crop modelling, such as accounting for evolving [CO_2]. The inclusion of elevated [CO_2] is of vital importance due to the unrealistic nature of fixing a baseline [CO_2] level for future time periods and the known importance of [CO_2] on growth of C_3 crops such as barley and wheat (Drake et al., 1997). Although uncertainty remains in the magnitude of the [CO_2] effect (Soussana et al., 2010), the parameter values of the [CO_2] response used in the study are conservative, based on FACE experiments, showing lower responses than earlier pot and chamber experiments (Long et al., 2006; Weigel and Manderscheid, 2012). It should be noted that in papers **I** and **II** the primary interest was on the sensitivity of simulated yields to systematic changes in temperature and precipitation across a large ensemble, and thus other variables were fixed on purpose.

The temperature-based sowing date estimation method was implemented in paper **III** as an improvement in realism to applying a fixed sowing date across years/changes to climate, but also as a means to evaluate the effectiveness of changing the time of sowing as an adaptation option. The estimated dates are in line with results from other research for similar scenarios of future climate change (e.g. Olesen et al., 2012; Peltonen-Sainio et al., 2009a) and the method is easy to apply. Nonetheless, better availability of geo-located paired sowing date and temperature observations would enable further verification of results and allow possible refinements of the chosen approach, e.g. suited for use in sub-regions across Finland.

The decision made on applying a harvest cut-off was found to be essential for treating simulations that reported yields with maturity/harvest dates beyond plausible windows for harvest. The effect of applying the dynamically defined harvest cut-off (paper **III**) as opposed to a fixed date (papers **I** and **II)** made only very little difference in Finland. However, the temperature-based definition gives more flexibility with respect to yearly variation and thus can be seen as a potential improvement in its application. The decision of setting yields as a result to zero remains debatable and raises wider questions concerning the treatment of model output data. Alternative approaches could involve using reduction factors to progressively reduce the yield from an initial cut-off towards an absolute one. However, observational data (not available in the context of this study) would be needed to support the development of such methods. Setting yields abruptly to zero has implications on the distribution of yields across the 30-year period affecting, for example, period averages with mean and median producing quite different results. Further, interpolation of yields on the IRS between yield values and truncated values of zero causes an unrealistically rapid decline from one temperature change to another. Consequently, results for cooling, where the cut-off occasionally applies, need to be treated with caution.

4.6 OUTLOOK

The analysis and outcomes of this study highlight several requirements and opportunities for future work. The study demonstrated aspects of applying different multi-model ensembles individually with respect to crop models in papers **I** and **II,** and to climate in paper **III.** By combining these different ensemble approaches to estimate impact likelihoods, using the IRS-based approach, both climate and impact model uncertainties could be addressed. Examples that apply such combinations of ensembles have been presented e.g. for palsa mires by Fronzek et al. (2011) and crops by Tao et al. (2018), both applying an ensemble of models, multiple sets of model parameters and a set of contrasting climate projections together to conduct a probabilistic assessment of climate change impacts. The idea of applying multi-model ensembles could further be extended to using an “ensemble of users" for each model, each applying their own methods of model calibration. In addition to an overall need to pay more attention to the consistency of model calibration, this could offer a way of accounting for calibration uncertainty in impact estimates, which is rarely implemented in impact studies (Confalonieri et al., 2016). Furthermore, the methods developed in this study for classifying IRSs could be applied to other regions/crops/sectors. They could then be used to inform model selection in future multi-model simulation studies through identification of model responses that exhibit similarities as well as outliers clearly exhibiting divergent behaviour.

Following suggestions by Holman et al. (2019) on improving simulations of adaptation, there are already plans to work in close collaboration with stakeholders to assess the choice of relevant indicators of climate change impacts on agriculture and to relate likelihood estimations to meaningful, critical thresholds. The aim is to understand better the triggers and goals of adaptation in order to feed into on-going policy efforts for improving the resilience of crop production under future climate change. Meaningful adaptation planning should also include a more comprehensive exploration of available options and inter-linkages between them including considerations of farm interactions, crop rotations and constraints posed by available resources and other real-world capacities. By focusing on crop production alone, the influence of other factors on cereal yields through changes in demand, land suitability and resource competition is omitted, and the magnitude of impacts possibly misrepresented. The same applies to using crop models that fail to account for stresses due to issues such as weeds, pests, diseases and extreme weather events, occurrences of which are likely to change in the future. More broadly, in aiming towards informing adaptation policy, the importance of undertaking integrated, cross-sectoral assessments of climate change impacts has also been highlighted (e.g. Harrison et al., 2016).

For supporting efforts in developing regional and national adaptation plans and feeding into national risk assessments, new work can extend the analysis of yield likelihoods geographically from a single site to the regional level through a set of representative sites or across a regular grid. The scope of studies should also be widened from a single sector focus to looking at likelihoods across sectors. Estimates of evolving impact likelihoods can be compared in a way that assesses urgency for action, sharing characteristics with the reasons for concern used in the IPCC assessments (IPCC, 2014b). This could involve, for example, colour coded mapping of impact likelihoods that indicates the level of risk at a regional level.

5 CONCLUSIONS

This thesis has demonstrated the application of an IRS-based approach in a risk framework for estimating the sensitivity of cereal yields to perturbed climate, for computing future likelihoods of yield shortfall using projections of climate interpreted probabilistically and for evaluating the potential effectiveness of adaptation measures in improving crop performance under changing conditions. The visual representation of simulation results allows for an improved appreciation of different outcomes to aid efforts at explaining them, assisting future efforts to improve both the models themselves and the methods used to apply them. While previous studies have explored various aspects presented here, this is the first to focus systematically on addressing crop model sensitivities in a multi-model framework across sites in Europe and also the first to apply the IRS approach to examine crop yield likelihoods with and without adaptation. Novel methods for classifying patterns of impact response were also developed. The main conclusions of this study are as follows, grouped according to the specific objectives of the thesis:

Performance of crop models and sensitivities of cereal yields under present-day conditions and assumed changes in climate (Objective 1)

- Large variation across individual models was found in the simulation of year-to-year yields under the baseline climate. Comparison of observed vs simulated yields reflected a yield gap often associated with such comparisons. Proper evaluation of model performance was hampered by a lack of consistent observed crop data, which highlights an important need for improved monitoring and data collection.
- Across sites, simulated cereal yields declined with warming and drying but benefited from increased precipitation.
- The sensitivities of simulated cereal yield varied across the latitudinal transect of sites in Finland, Germany and Spain. At the Finnish site multi-model median wheat yields are more sensitive to changes in temperature than precipitation. In Germany and Spain sensitivities are more evenly distributed among the two drivers for winter wheat. Response of spring wheat in Spain is precipitation dominated.
- Under future scenarios, elevated [CO_2] compensated for the negative impacts of higher temperatures simulated for barley yields in Finland, with the impact strongest at the largest temperature changes.

Differences in model behaviour across an ensemble of wheat models in different parts of Europe (Objective 2)

- Two new, complementary methods for classifying and interpreting IRS patterns have been developed, an expert and a statistical diagnostic approach (EDA and SDA).
- Similarity in the pattern of modelled yield sensitivity is greatest for Germany across a multi-model ensemble, in comparison to Finland and Spain. This suggests that divergence in model behaviour occurs when attempting to represent stresses affecting yields that are found more often in Finland and Spain. These include heat and drought stress, effects of cold and treatment of vernalisation for winter varieties.
- Greater similarity in responses was found for models belonging to the same model family than for those not characterised as belonging to any. No clear messages were found with respect to specific properties of the models that explain the variation in the impact responses across the models.

Likelihood of future yield impacts on barley in Finland (Objective 3)

- By combining IRSs with projections of climate change interpreted probabilistically, with and without the effect of increasing future [CO_2], the evolution of likelihood of barley yield shortfall throughout the century was analysed and visualised.
- The yield losses associated with warming are compensated with the beneficial effects of evolving [CO_2] levels on yields. The degree of warming relative to the increase in [CO_2] was found to show in the results of evolving likelihoods on yield impacts, with RCP8.5 benefitting yields more than RCP4.5 due to its higher level of [CO_2] relative to warming.

Effectiveness of potential farm-level adaptation options on barley in Finland (Objective 4)

- Adapting sowing to yearly weather conditions results in earlier sowing under projected warming and lowers the likelihood of yield shortfall due to better timing of critical growth phases with respect to seasonal growing conditions.
- Combining cultivars with short pre- and long post-anthesis phases with earlier sowing, is likely to benefit yields the most under projected climate change.

Implications of using different approaches for applying models and analysing their results (Objective 5)

- The IRS approach represents sensitivities of an impact variable to changes in two key drivers. By spanning a range of plausible outcomes, estimates of impacts under any future projection of the same two variables used in its construction can be obtained without the need to conduct new simulations. Hence, it is "scenario-neutral".
- The IRS approach can assist in: (i) identifying possible needs for model development through identification of patterns and irregularities of model behaviour, (ii) comparison of sensitivities across models, studies and/or sectors and for different options such as choice of crop or soil, (iii) examination of various statistical characteristics of the response across a large range of changes to key drivers, including absolute mean yield, change relative to baseline yield, inter-model and inter-annual spread.
- The choice of impact model and how the model and study are set-up can have a large effect on the results. Key considerations include inclusion of [CO_2], the choice and representation of soils and the method of calibration.
- Individual models represent different degrees of detail in simulating different processes. While no model can fully represent all relevant processes, where the emphasis has been placed in a particular model can aid in selecting a single model for the question at hand.
- Use of a multi-model ensemble for both climate projections and the modelling of crop responses, as opposed to relying on the outputs of a single model, increases the robustness of the results and provides information on the uncertainty around the yield estimates.
- In comparison to conventional and more detailed scenario-based approaches, and when accounting for seasonality in changes of temperature and precipitation, the IRS approach offers a valid tool for examining yield impacts and associated likelihoods.

REFERENCES

Ainsworth, E.A. and Long, S.P., 2005. What have we learned from 15 years of free-air CO_2 enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO_2. New Phytologist, 165(2): 351-372.

Allen Jr., L.H., 2019. Simplifying Crop Growth Response to Rising CO_2 and Elevated Temperature. Agricultural & Environmental Letters, 4(1): 190021.

Allen, R.G., Pereira, L.S., Raes, D. and Smith, M., 1998. Crop evapotranspiration - Guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper No. 56, Rome, Italy, 300 pp.

Amthor, J.S., 2001. Effects of atmospheric CO_2 concentration on wheat yield: review of results from experiments using various approaches to control CO_2 concentration. Field Crops Research, 73(1): 1-34.

Angulo, C., Rötter, R., Trnka, M., Pirttioja, N., Gaiser, T., Hlavinka, P. and Ewert, F., 2013. Characteristic 'fingerprints' of crop model responses to weather input data at different spatial resolutions. European Journal of Agronomy, 49: 104-114.

Asseng, S., Ewert, F., Martre, P., Rötter, R.P., Lobell, D.B., Cammarano, D., Kimball, B.A., Ottman, M.J., Wall, G.W., White, J.W. and 43 others, 2014a. Rising temperatures reduce global wheat production. Nature Climate Change.

Asseng, S., Ewert, F., Rosenzweig, C., Jones, J.W., Hatfield, J.L., Ruane, A.C., Boote, K.J., Thorburn, P.J., Rötter, R.P., Cammarano, D. and 41 others, 2013. Uncertainty in simulating wheat yields under climate change. Nature Climate Change, 3: 827-832.

Asseng, S., Zhu, Y., Basso, B., Wilson, T. and Cammarano, D., 2014b. Simulation Modeling: Applications in Cropping Systems. In: N.K. Van Alfen (Editor), Encyclopedia of Agriculture and Food Systems. Academic Press, Oxford, pp. 102-112.

Bassu, S., Brisson, N., Durand, J.-L., Boote, K., Lizaso, J., Jones, J.W., Rosenzweig, C., Ruane, A.C., Adam, M., Baron, C. and 28 others, 2014. How do various maize crop models vary in their responses to climate change factors? Global Change Biology, 20(7): 2301-2320.

Beillouin, D., Schauberger, B., Bastos, A., Ciais, P. and Makowski, D., 2020. Impact of extreme weather conditions on European crop production in 2018. Philosophical Transactions of the Royal Society B: Biological Sciences, 375(1810): 20190510.

Beven, K. and Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. J. Hydrol., 249(1): 11-29.

Bindoff, N., Stott, P., AchutaRao, K.M., Allen, M., Gillett, N., Gutzler, D., Hansingo, K., Hegerl, G., Hu, Y., Jain, S. and 5 others, 2013. Detection and attribution of climate change: from global to regional. In: T.F. Stocker et al. (Editors), Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 867-952.

Brisson, N., Gate, P., Gouache, D., Charmet, G., Oury, F.-X. and Huard, F., 2010. Why are wheat yields stagnating in Europe? A comprehensive data analysis for France. Field Crops Research, 119(1): 201-212.

Børgesen, C.D. and Olesen, J.E., 2011. A probabilistic assessment of climate change impacts on yield and nitrogen leaching from winter wheat in Denmark. Natural Hazards and Earth System Science, 11(9): 2541-2553.

Challinor, A., Watson, J., Lobell, D., Howden, S., Smith, D. and Chhetri, N., 2014. A meta-analysis of crop yield under climate change and adaptation. Nature Climate Change, 4(4): 287-291.

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., Gao, X., Gutowski, W., Johns, T., Krinner, G. et al., 2013. Long-term climate change: projections, commitments and irreversibility. In: T.F. Stocker et al. (Editors), Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1029-1136, pp. 1029-1136.

Confalonieri, R., Orlando, F., Paleari, L., Stella, T., Gilardelli, C., Movedi, E., Pagani, V., Cappelli, G., Vertemara, A., Alberti, L. and 26 others, 2016. Uncertainty in crop model predictions: What is the role of users? Environmental Modelling & Software, 81: 165-173.

Crespo-Herrera, L.A., Crossa, J., Huerta-Espino, J., Vargas, M., Mondal, S., Velu, G., Payne, T.S., Braun, H. and Singh, R.P., 2018. Genetic Gains for Grain Yield in CIMMYT's Semi-Arid Wheat Yield Trials Grown in Suboptimal Environments. Crop Science, 58(5): 1890-1898.

Cubasch, U., Wuebbles, D., Chen, D., Facchini, M.C., Frame, D., Mahowald, N. and Winther, J.-G., 2013. Introduction. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. In: T.F. Stocker et al. (Editors). Cambridge University Press,, Cambridge, United Kingdom and New York, NY, USA.

de Wit, A., Boogaard, H., Fumagalli, D., Janssen, S., Knapen, R., van Kraalingen, D., Supit, I., van der Wijngaart, R. and van Diepen, K., 2019. 25 years of the WOFOST cropping system model. Agric. Syst., 168: 154-167.

de Wit, C.T., 1958. Transpiration and crop yields, Vol. 64.6. Verslagen van Landbouwkundige Onderzoekingen, Institute of Biological and Chemical Research on Field Crops and Herbage. 0372-6223, Wageningen, The Netherlands.

de Wit, C.T., 1965. Photosynthesis of leaf canopies, Pudoc, Wageningen.

de Wit, C.T., 1982. Simulation of living systems. In: F.W.T. Penning de Vries and H.H. van Laar (Editors), Simulation of plant growth and crop production. Centre for Agricultural Publishing and Documentation (Pudoc), Wageningen.

Diaz-Nieto, J. and Wilby, R.L., 2005. A comparison of statistical downscaling and climate change factor methods: impacts on low flows in the River Thames, United Kingdom. Climatic Change, 69(2-3): 245-268.

Donatelli, M., Magarey, R.D., Bregaglio, S., Willocquet, L., Whish, J.P.M. and Savary, S., 2017. Modelling the impacts of pests and diseases on agricultural systems. Agric. Syst., 155: 213-224.

Drake, B.G., Gonzàlez-Meler, M.A. and Long, S.P., 1997. More Efficient Plants: A Consequence of Rising Atmospheric CO_2? Annual Review of Plant Physiology and Plant Molecular Biology, 48(1): 609-639.

Duncan, W.G., Loomis, R.S., Williams, W.A. and Hanau, R., 1967. A model for simulating photosynthesis in plant communities. Hilgardia, 38(4): 181-205.

Easterling, W.E., Chen, X., Hays, C., Brandle, J.R. and Zhang, H., 1996. Improving the validation of model-simulated crop yield response to climate change: an application to the EPIC model. Climate Research, 6(3): 263-273.

Easterling, W.E., McKenney, M.S., Rosenberg, N.J. and Lemon, K.M., 1992. Simulations of crop response to climate change: effects with present technology and no adjustments (the 'dumb farmer' scenario). Agricultural and Forest Meteorology, 59(1): 53-73.

European Commision, 2019. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, The European Green Deal, COM/2019/640 final.

European Commision, 2021. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Forging a climate-reslient Europe - the new EU strategy on Adpatation to Climate Change, COM/2021/82 final.

EUROSTAT, 2014. Crop products yields by NUTS 2 regions, http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database.

Evans, L.T., 1996. Crop evolution, adaptation and yield. Cambridge university press.

Evenson, R.E. and Gollin, D., 2003. Assessing the Impact of the Green Revolution, 1960 to 2000. Science, 300(5620): 758-762.

Ewert, F., Rötter, R.P., Bindi, M., Webber, H., Trnka, M., Kersebaum, K.C., Olesen, J.E., van Ittersum, M.K., Janssen, S., Rivington, M. and 12 others, 2015. Crop modelling for integrated assessment of risk to food production from climate change. Environmental Modelling & Software, 72: 287-303.

Falloon, P., Challinor, A., Dessai, S., Hoang, L., Johnson, J. and Koehler, A.-K., 2014. Ensembles and uncertainty in climate change impacts. Frontiers in Environmental Science, 2(33).

FAO, 2002. Bread Wheat Improvement and production. FAO Plant Production and Protection Series, 30, Rome. http://www.fao.org/3/y4011e/y4011e00.htm#Contents.

FAOSTAT, 2018. Crops, http://www.fao.org/faostat/en/#data/QC.

Farquhar, G.D., von Caemmerer, S. and Berry, J.A., 1980. A biochemical model of photosynthetic CO_2 assimilation in leaves of C3 species. Planta, 149: 78-90.

Ferrise, R., Moriondo, M. and Bindi, M., 2011. Probabilistic assessments of climate change impacts on durum wheat in the Mediterranean region. Natural Hazards and Earth System Sciences, 11(5): 1293-1302.

Finnish Cereal Committee, 2014. Production of cereal and oilseed crops in Finland. https://www.vyr.fi/document/1/137/250801f/muutra_217123e_Vilja_alanyhteisty_esite_eng_30_5_LOW.pdf.

Food and Agriculture Organization of the United Nations, 2017. FAOSTAT Database. FAO. http://www.fao.org/faostat/en/#data/QC, Rome, Italy.

Forkman, J., Amiri, S. and von Rosen, D., 2012. Effect of region on the uncertainty in crop variety trial programs with a reduced number of trials. Euphytica, 186(2): 489-500.

Fronzek, S., 2013. Climate change and the future distribution of palsa mires: ensemble modelling, probabilities and uncertainties. Monographs of the Boreal Environmental Research No. 44, ISBN 978-952-11-4204-8, 35 pp.

Fronzek, S., Carter, T.R. and Luoto, M., 2011. Evaluating sources of uncertainty in modelling the impact of probabilistic climate change on sub-arctic palsa mires. Natural Hazards and Earth System Sciences, 11: 2981-2995.

Fronzek, S., Carter, T.R., Pirttioja, N., Alkemade, R., Audsley, E., Bugmann, H., Flörke, M., Holman, I., Honda, Y., Ito, A. and 11 others, 2018. Determining sectoral and regional sensitivity to climate and socio-economic change in Europe using impact response surfaces. Regional Environmental Change.

Fronzek, S., Carter, T.R., Raisanen, J., Ruokolainen, L. and Luoto, M., 2010. Applying probabilistic projections of climate change with impact models: a case study for sub-arctic palsa mires in Fennoscandia. Climatic Change, 99(3-4): 515-534.

Füssel, H.-M., Jol, A., Marx, A., Hildén, M., Aparicio, A., Bastrup-Birk, A., Bigano, A., Castellari, S., Erhard, M., Georgi, B. and 59 others, 2017. Climate change, impacts and vulnerability in Europe 2016 - An indicator-based report, 1/2017. Publications Office of the European Union, 419 pp.

Gay, C. and Estrada, F., 2010. Objective probabilities about future climate are a matter of opinion. Climatic Change, 99(1): 27-46.

Giorgi, F., 2005. Climate Change Prediction. Climatic Change, 73(3): 239-265.

Grassini, P., van Bussel, L.G.J., Van Wart, J., Wolf, J., Claessens, L., Yang, H., Boogaard, H., de Groot, H., van Ittersum, M.K. and Cassman, K.G., 2015. How good is good enough? Data requirements for reliable crop yield simulations and yield-gap analysis. Field Crops Research, 177: 49-63.

Harris, G.R., Collins, M., Sexton, D.M.H., Murphy, J.M. and Booth, B.B.B., 2010. Probabilistic projections for 21st century European climate. Natural Hazards and Earth System Sciences, 10(9): 2009-2020.

Harrison, P.A., Dunford, R.W., Holman, I.P. and Rounsevell, M.D.A., 2016. Climate change impact modelling needs to include cross-sectoral interactions. Nature Climate Change, 6(9): 885-890.

Hartmann, D.L., Klein Tank, A.M.G., Rusticucci, M., Alexander, L.V., Brönnimann, S., Charabi, Y., Dentener, F.J., Dlugokencky, E.J., Easterling, D.R., Kaplan, A. and 4 others, 2013. Observations: Atmosphere and Surface. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. In: T.F. Stocker et al. (Editors). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Hatfield, J.L. and Prueger, J.H., 2015. Temperature extremes: Effect on plant growth and development. Weather and Climate Extremes, 10: 4-10.

He, J., Jones, J.W., Graham, W.D. and Dukes, M.D., 2010. Influence of likelihood function choice for estimating crop model parameters using the generalized likelihood uncertainty estimation method. Agric. Syst., 103(5): 256-264.

Hinkel, J., Bharwani, S., Bisaro, A., Carter, T., Cull, T., Davis, M., Klein, R., Lonsdale, K., Rosentrater, L. and Vincent, K., 2013. PROVIA Guidance on Assessing Vulnerability, Impacts and Adpatation to Climate Change, United Nations Environment Programme.

Hoegh-Guldberg, O., Jacob, D., Taylor, M., Bindi, M., Brown, S., Camilloni, I., Diedhiou, A., Djalante, R., Ebi, K.L., Engelbrecht, F. and 8 others, 2018. Impacts of 1.5ºC Global Warming on Natural and Human Systems. In: V. Masson-Delmotte et al. (Editors), Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty, In Press.

Holman, I.P., Brown, C., Carter, T.R., Harrison, P.A. and Rounsevell, M., 2019. Improving the representation of adaptation in climate change impact models. Regional Environmental Change, 19(3): 711-721.

Holmberg, M., Futter, M.N., Kotamäki, N., Fronzek, S., Forsius, M., Kiuru, P., Pirttioja, N., Rasmus, K., Starr, M. and Vuorenmaa, J., 2014. Effects of changing climate on the hydrology of a boreal catchment and lake DOC--probabilistic assessment of a dynamic model chain. Boreal Environment Research, 19: 66-82.

Howden, S.M., Soussana, J.-F., Tubiello, F.N., Chhetri, N., Dunlop, M. and Meinke, H., 2007. Adapting agriculture to climate change. Proceedings of the National Academy of Sciences, 104(50): 19691-19696.

Iglesias, A., Garrote, L., Quiroga, S. and Moneo, M., 2012. A regional comparison of the effects of climate change on agricultural crops in Europe. Climatic Change, 112(1): 29-46.

IPCC, 2013. Annex II: Climate System Scenario Tables [Prather M, Flato G, Friedlingstein P, Jones C, Lamarque J-F, Liao H, Rasch P (eds.)]. In: T. Stocker et al. (Editors), Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1395-1445.

IPCC, 2014a. Annex II: Glossary [Agard, J. et al. (eds.)]. In: V.R. Barros et al. (Editors), Climate Change 2014: Impacts, Adaptation, and Vulnerability Part B: Regional Aspects Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1757-1776.

IPCC, 2014b. Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. In: Core Writing Team, R.K. Pachauri and L.A. Meyer (Editors). IPCC, Geneva, Switzerland, pp. 151.

IPCC, 2014c. Summary for policymakers. In: C.B. Field et al. (Editors), Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1-32.

IPCC, 2018. Global Warming of 1.5 °C. An IPCC special report on the impacts of global warming of 1.5 °C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty. In: V. Masson-Delmotte et al. (Editors). In Press.

Jones, J.W., Antle, J.M., Basso, B., Boote, K.J., Conant, R.T., Foster, I., Godfray, H.C.J., Herrero, M., Howitt, R.E., Janssen, S. et al., 2017. Brief history of agricultural systems modeling. Agric. Syst., 155: 240-254.

Jones, R. and Mearns, L., 2005. Chapter 5: Assessing Future Climate Risks. In: B. Lim, E. Spanger-Siegfried, I. Burton, E. Malone and S. Huq (Editors), Adaptation Policy Frameworks for Climate Change: Devloping Strategies, Policies and Measures. Cambridge University Press, Cambridge, UK, pp. 119-144.

Jones, R.N., 2000. Managing uncertainty in climate change projections - Issues for impact assessment - An editorial comment. Climatic Change, 45(3-4): 403-419.

Kangas, A., Laine, A., Niskanen, M., Salo, Y., Vuorinen, M., Jauhiainen, L. and Nikander, H., 2010. Results of official variety trials (In Finnish). MTT Agrifood Research Finland, Jokioinen, Finland, 175 pp.

Kant, S., Seneweera, S., Rodin, J., Materne, M., Burch, D., Rothstein, S.J. and Spangenberg, G., 2012. Improving yield potential in crops under elevated CO_2: Integrating the photosynthetic and nitrogen utilization efficiencies. Frontiers in plant science, 3: 1-9.

Kaukoranta, T. and Hakala, K., 2008. Impact of spring warming on sowing times of cereal, potato and sugar beet in Finland. Agricultural and Food Science, 17(2): 165-176.

Kay, A.L., Crooks, S.M., Davies, H.N., Prudhomme, C. and Reynard, N.S., 2014. Probabilistic impacts of climate change on flood frequency using response surfaces I: England and Wales. Regional Environmental Change, 14(3): 1215-1227.

Keller, L., Rössler, O., Martius, O. and Weingartner, R., 2019. Comparison of scenario-neutral approaches for estimation of climate change impacts on flood characteristics. Hydrological Processes, 33(4): 535-550.

Knutti, R., Joos, F., Müller, S.A., Plattner, G.-K. and Stocker, T.F., 2005. Probabilistic climate change projections for CO_2 stabilization profiles. Geophys. Res. Lett., 32(20).

Kontturi, M., 1979. The effect of weather on yield and development of spring wheat in Finland. Annales Agriculturae Fenniae, 18: 263-274.

Koponen, J. and Hildén, J., 2019. Data visualization handbook. Otavan kirjapaino, Keuruu.

Kovats, R.S., Valentini, R., Bouwer, L.M., Georgopoulou, E., Jacob, D., Martin, E., Rounsevell, M. and Soussana, J.-F., 2014. Europe. In: V.R. Barros et al. (Editors), Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1267-1326.

Kristensen, K., Schelde, K. and Olesen, J.E., 2011. Winter wheat yield response to climate variability in Denmark. The Journal of Agricultural Science, 149(1): 33-47.

Kroes, J.G., Van Dam, J.C., Groenendijk, P., Hendriks, R.F.A. and Jacobs, C.M.J., 2008. SWAP Version 4: Theory Description and User Manual. Wageningen Environmental Research, Wageningen. http://edepot.wur.nl/416321C1.

Kroes, J.G., Wesseling, J.G. and Van Dam, J.C., 2000. Integrated modelling of the soil–water–atmosphere–plant system using the model SWAP 2·0 an overview of theory and an application. Hydrological Processes, 14(11-12): 1993-2002.

Lachaut, T. and Tilmant, A., 2020. Possibilistic response surfaces combining fuzzy targets and hydro-climatic uncertainty in flood vulnerability assessment. Hydrol. Earth Syst. Sci. Discuss., 2020: 1-30.

Lal, P.N., Mitchell, T., Aldunce, P., Auld, H., Mechler, R., Miyan, A., Romano, L.E. and Zakaria, S., 2012. National systems for managing the risks from climate extremes and disasters. In: C.B. Field et al. (Editors), A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC). Cambridge University Press, Cambridge, UK and New York, NY, USA, pp. 339-392.

Lavell, A., Oppenheimer, M., Diop, C., Hess, J., Lempert, R., Li, J., Muir-Wood, R. and Myeong, S., 2012. Climate change: new dimensions in disaster risk, exposure, vulnerability, and resilience. In: C.B. Field et al. (Editors), Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC). Cambridge University Press, Cambridge, YK and New York, NY, USA, pp. 25-64.

Lawlor, D.W. and Mitchell, R.A.C., 2000. Crop Ecosystem Responses to Climatic Change: Wheat. In: K.R. Reddy and H.F. Hodges (Editors), Climate Change and Global Productivity. CAB International, New York, pp. 57-80.

Levis, S., Badger, A., Drewniak, B., Nevison, C. and Ren, X., 2018. CLMcrop yields and water requirements: avoided impacts by choosing RCP 4.5 over 8.5. Climatic Change, 146(3): 501-515.

Li, Z.T., Yang, J.Y., Smith, W.N., Drury, C.F., Lemke, R.L., Grant, B., He, W.T. and Li, X.G., 2015. Simulation of long-term spring wheat yields, soil organic C, N and water dynamics using DSSAT-CSM in a semi-arid region of the Canadian prairies. Nutrient Cycling in Agroecosystems, 101(3): 401-419.

Lipper, L., Thornton, P., Campbell, B.M., Baedeker, T., Braimoh, A., Bwalya, M., Caron, P., Cattaneo, A., Garrity, D., Henry, K. and 14 others, 2014. Climate-smart agriculture for food security. Nature Climate Change, 4: 1068-1072.

Liu, B., Asseng, S., Liu, L., Tang, L., Cao, W. and Zhu, Y., 2016. Testing the responses of four wheat crop models to heat stress at anthesis and grain filling. Glob Chang Biol, 22(5): 1890-903.

Liu, Y., Chen, Q., Ge, Q., Dai, J. and Dou, Y., 2018. Effects of climate change and agronomic practice on changes in wheat phenology. Climatic Change, 150(3): 273-287.

Lobell, D.B. and Burke, M.B., 2010. On the use of statistical models to predict crop yield responses to climate change. Agricultural and Forest Meteorology, 150(11): 1443-1452.

Lobell, D.B. and Field, C.B., 2007. Global scale climate–crop yield relationships and the impacts of recent warming. Environmental research letters, 2(1).

Long, S.P., Ainsworth, E.A., Leakey, A.D., Nösberger, J. and Ort, D.R., 2006. Food for thought: lower-than-expected crop yield stimulation with rising CO_2 concentrations. Science, 312(5782): 1918-1921.

Lorenz, D.J. and DeWeaver, E.T., 2007. The response of the extratropical hydrological cycle to global warming. Journal of climate, 20(14): 3470-3484.

Luo, Q., Bellotti, W., Williams, M. and Bryan, B., 2005. Potential impact of climate change on wheat yield in South Australia. Agricultural and Forest Meteorology, 132(3): 273-285.

Luo, Q., Bellotti, W., Williams, M., Cooper, I. and Bryan, B., 2007. Risk analysis of possible impacts of climate change on South Australian wheat production. Climatic Change, 85(1-2): 89-101.

MAGRAMA, 2010. Anuario de Estadística. Ministerio de Agricultura, Alimentación y Medio Ambiente, Madrid (MAGRAMA), Spain.

Maiorano, A., Martre, P., Asseng, S., Ewert, F., Müller, C., Rötter, R.P., Ruane, A.C., Semenov, M.A., Wallach, D., Wang, E. and 27 others, 2017. Crop model improvement reduces the uncertainty of the response to temperature of multi-model ensembles. Field Crops Research, 202: 5-20.

Martre, P., Wallach, D., Asseng, S., Ewert, F., Jones, J.W., Rötter, R.P., Boote, K.J., Ruane, A.C., Thorburn, P.J., Cammarano, D. and 40 others, 2015. Multimodel ensembles of wheat growth: Many models are better than one. Global change biology, 21: 911-925.

Metzger, M., Bunce, R., Jongman, R., Mücher, C. and Watkins, J., 2005. A climatic stratification of the environment of Europe. Global ecology and biogeography, 14(6): 549-563.

Micale, F. and Genovese, G., 2004. Methodology of the MARS crop yield forecasting system. Vol. 1: meteorological data collection, processing and analysis. Contributions from: E. Van der Goot, I. Supit, H. Boogard, K. Van Diepen, F. Micale, S. Orlandi, H. Otten, M. Geuze, D. Schulze. EUR, 21291.

Mínguez, M.I., Ruiz-Ramos, M., Díaz-Ambrona, C.H., Quemada, M. and Sau, F., 2007. First-order impacts on winter and summer crops assessed with various high-resolution climate models in the Iberian Peninsula. Climatic Change, 81(1): 343-355.

Ministerio de Agricultura, Pesca y Alimentación, 2018. Estadísticas agrarias: Agricultura, www.mapa.gob.es.

MMM, 2014. Finland's National Climate Change Adaptation Plan 2022. Publications of the Ministry of Agriculture and Forestry 5b/2014:40 pp.

Monfreda, C., Ramankutty, N. and Foley, J.A., 2008. Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. Global biogeochemical cycles, 22(1).

Monteith, J.L., 1981. Climatic variation and the growth of crops. Quarterly Journal of the Royal Meteorological Society, 107: 749-774.

Moore, F.C. and Lobell, D.B., 2015. The fingerprint of climate trends on European crop yields. Proceedings of the National Academy of Sciences, 112(9): 2670-2675.

Morison, J.I.L., Baker, N.R., Mullineaux, P.M. and Davies, W.J., 2008. Improving water use in crop production. Philosophical Transactions of the Royal Society B: Biological Sciences, 363(1491): 639-658.

Murphy, J.M., Sexton, D.M., Jenkins, G.J., Booth, B.B., Brown, C.C., Clark, R.T., Collins, M., Harris, G.R., Kendon, E.J., Betts, R.A. and 9 others, 2009. UK Climate Projections Science Report: Climate Change projections. Met Office Hadley Centre, Exeter.

Mäkelä, H.M., Venäläinen, A., Jylhä, K., Lehtonen, I. and Gregow, H., 2014. Probabilistic projections of climatological forest fire danger in Finland. Climate Research, 60(1): 73-85.

Noble, I.R., Huq, S., Anokhin, Y.U., Carmin, J., Goudou, D., Lansigan, F.P., Osman-Elasha, B. and Villamizar, A., 2014. Adaptation Needs and Options. In: C.B. Field et al. (Editors), Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 833-868.

O'Neill, B.C., Tebaldi, C., van Vuuren, D.P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.F., Lowe, J. and 4 others, 2016. The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. Geosci. Model Dev., 9(9): 3461-3482.

Oerke, E.-C., 2006. Crop losses to pests. Journal of Agricultural Science, 144: 31-43.

Olesen, J.E. and Bindi, M., 2002. Consequences of climate change for European agricultural productivity, land use and policy. European Journal of Agronomy, 16(4): 239-262.

Olesen, J.E., Børgesen, C.D., Elsgaard, L., Palosuo, T., Rötter, R.P., Skjelvåg, A.O., Peltonen-Sainio, P., Börjesson, T., Trnka, M., Ewert, F. and 6 others, 2012. Changes in time of sowing, flowering and maturity of cereals in Europe under climate change. Food Additives & Contaminants: Part A, 29(10): 1527-1542.

Olesen, J.E., Carter, T.R., Díaz-Ambrona, C.H., Fronzek, S., Heidmann, T., Hickler, T., Holt, T., Minguez, M.I., Morales, P., Palutikof, J.P. and 6 others, 2007. Uncertainties in projected impacts of climate change on European agriculture and terrestrial ecosystems based on scenarios from regional climate models. Climatic Change, 81(1): 123-143.

Olesen, J.E., Trnka, M., Kersebaum, K.C., Skjelvåg, A.O., Seguin, B., Peltonen-Sainio, P., Rossi, F., Kozyra, J. and Micale, F., 2011. Impacts and adaptation of European crop production systems to climate change. European Journal of Agronomy, 34(2): 96-112.

Palosuo, T., Kersebaum, K.C., Angulo, C., Hlavinka, P., Moriondo, M., Olesen, J.E., Patil, R.H., Ruget, F., Rumbaur, C., Takáč, J. and 9 others, 2011. Simulation of winter wheat yield and its variability in different climates of Europe: A comparison of eight crop growth models. European Journal of Agronomy 35: 103-114.

Palosuo, T., Rötter, R.P., Salo, T., Peltonen-Sainio, P., Tao, F. and Lehtonen, H., 2015. Effects of climate and historical adaptation measures on barley yield trends in Finland. Climate Research, 65: 221-236.

Panda, A., 2018. Transformational adaptation of agricultural systems to climate change. Wiley Interdisciplinary Reviews: Climate Change, 9(4).

Peltonen-Sainio, P. and Jauhiainen, L., 2020. Large zonal and temporal shifts in crops and cultivars coincide with warmer growing seasons in Finland. Regional Environmental Change, 20(3): 89.

Peltonen-Sainio, P., Jauhiainen, L. and Hakala, K., 2011. Crop responses to temperature and precipitation according to long-term multi-location trials at high-latitude conditions. The Journal of Agricultural Science, 149(01): 49-62.

Peltonen-Sainio, P., Jauhiainen, L., Hakala, K. and Ojanen, H., 2009a. Climate change and prolongation of growing season: changes in regional potential for field crop production in Finland. Agricultural and Food Science, 18: 171-190.

Peltonen-Sainio, P., Jauhiainen, L. and Laurila, I.P., 2009b. Cereal yield trends in northern European conditions: Changes in yield potential and its realisation. Field Crops Research, 110(1): 85-90.

Peltonen-Sainio, P., Palosuo, T., Ruosteenoja, K., Jauhiainen, L. and Ojanen, H., 2018. Warming autumns at high latitudes of Europe: an opportunity to lose or gain in cereal production? Regional Environmental Change.

Peltonen-Sainio, P. and Rajala, A., 2007. Duration of vegetative and generative development phases in oat cultivars released since 1921. Field Crops Research, 101(1): 72-79.

Porter, C.H., Villalobos, C., Holzworth, D., Nelson, R., White, J.W., Athanasiadis, I.N., Janssen, S., Ripoche, D., Cufi, J., Raes, D. and 5 others, 2014a. Harmonization and translation of crop modeling data to ensure interoperability. Environmental Modelling & Software, 62: 495-508.

Porter, J.R. and Gawith, M., 1999. Temperatures and the growth and development of wheat: a review. European Journal of Agronomy, 10: 23-26.

Porter, J.R., Xie, L., Challinor, A., Cochrane, K., Howden, M., Iqbal, M.M., Lobell, D.B. and Travasso, M.I., 2014b. Food security and food production systems. In: C.B. Field, V.R. Barros, M.D. Mastrandrea and K.J. Mach (Editors), Climate Change 2014: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 485-533.

Prudhomme, C., Wilby, R.L., Crooks, S., Kay, A.L. and Reynard, N.S., 2010. Scenario-neutral approach to climate change impact studies: Application to flood risk. J. Hydrol., 390(3): 198-209.

Ramirez-Villegas, J., Koehler, A.-K. and Challinor, A.J., 2017. Assessing uncertainty and complexity in regional-scale crop model simulations. European Journal of Agronomy, 88: 84-95.

Rembold, F., Meroni, M., Urbano, F., Csak, G., Kerdiles, H., Perez-Hoyos, A., Lemoine, G., Leo, O. and Negre, T., 2019. ASAP: A new global early warning system to detect anomaly hot spots of agricultural production for food security analysis. Agric. Syst., 168: 247-257.

Richards, L.A., 1931. Capillary conduction of liquids through porous mediums. Physics, 1(5): 318-333.

Rivington, M. and Koo, J., 2010. Report on the Meta-Analysis of Crop Modelling for Climate Change and Food Security Survey. CGIAR Program on Climate Change, Agriculture and Food Security (CCAFS), Copenhagen, Denmark.

Rodríguez, A., Ruiz-Ramos, M., Palosuo, T., Carter, T.R., Fronzek, S., Lorite, I.J., Ferrise, R., Pirttioja, N., Bindi, M., Baranowski, P. and 27 others, 2019. Implications of crop model ensemble size and composition for estimates of adaptation effects and agreement of recommendations. Agricultural and Forest Meteorology, 264: 351-362.

Rosenzweig, C. and Hillel, D., 1998. Climate Change and the Global Harvest. Oxford University Press, New York, 324 pp.

Rosenzweig, C., Iglesias, A., Yang, X.B., Epstein, P.R. and Chivian, E., 2001. Climate change and extreme weather events; implications for food production, plant diseases, and pests. Global change and human health, 2: 90-104.

Rosenzweig, C., Jones, J.W., Hatfield, J.L., Ruane, A.C., Boote, K.J., Thorburn, P., Antle, J.M., Nelson, G.C., Porter, C., Janssen, S. and 6 others, 2013. The Agricultural Model Intercomparison and Improvement Project (AgMIP): Protocols and pilot studies. Agricultural and Forest Meteorology, 170: 166-182.

Rosenzweig, C. and Parry, M.L., 1994. Potential impact of climate change on world food supply. Nature, 367(6459): 133-138.

Rosenzweig, C., Phillips, J., Goldberg, R., Carroll, J. and Hodges, T., 1996. Potential impacts of climate change on citrus and potato production in the US. Agric. Syst., 52(4): 455-479.

Ruane, A.C., McDermid, S., Rosenzweig, C., Baigorria, G.A., Jones, J.W., Romero, C.C. and DeWayne Cecil, L., 2014. Carbon–Temperature–Water change analysis for peanut production under climate change: a prototype for the AgMIP Coordinated Climate-Crop Modeling Project (C3MP). Global Change Biology, 20(2): 394-407.

Ruiz-Ramos, M., Ferrise, R., Rodríguez, A., Lorite, I.J., Bindi, M., Carter, T.R., Fronzek, S., Palosuo, T., Pirttioja, N., Baranowski, P. and 27 others, 2018. Adaptation response surfaces for managing wheat under perturbed climate and CO_2 in a Mediterranean environment. Agric. Syst., 159: 260-274.

Räisänen, J. and Palmer, T.N., 2001. A Probability and Decision-Model Analysis of a Multimodel Ensemble of Climate Change Simulations. Journal of Climate, 14(15): 3212-3226.

Räisänen, J. and Ruokolainen, L., 2006. Probabilistic forecasts of near-term climate change based on a resampling ensemble technique. Tellus A, 58(4): 461-472.

Rötter, R., Hoffmann, M., Koch, M. and Müller, C., 2018a. Progress in modelling agricultural impacts of and adaptations to climate change. Current Opinion in Plant Biology, 45: 255 - 261.

Rötter, R.P., Carter, T.R., Olesen, J.E. and Porter, J.R., 2011a. Crop-climate models need an overhaul. Nature Climate Change, 1(4): 175-177.

Rötter, R.P., Hoffmann, M.P., Koch, M. and Müller, C., 2018b. Progress in modelling agricultural impacts of and adaptations to climate change. Current Opinion in Plant Biology, 45: 1-7.

Rötter, R.P., Palosuo, T., Kersebaum, K.C., Angulo, C., Bindi, M., Ewert, F., Ferrise, R., Hlavinka, P., Moriondo, M. and Nendel, C., 2012. Simulation of spring barley yield in different climatic zones of Northern and Central Europe: a comparison of nine crop models. Field Crops Research, 133: 23-36.

Rötter, R.P., Palosuo, T., Pirttioja, N.K., Dubrovsky, M., Salo, T., Fronzek, S., Aikasalo, R., Trnka, M., Ristolainen, A. and Carter, T.R., 2011b. What would happen to barley production in Finland if global warming exceeded 4 degrees C? A model-based assessment. European Journal of Agronomy, 35(4): 205-214.

Sadras, V.O., Villalobos, F.J. and Fereres, E., 2016. Crop Development and Grwoth. In: F.J. Villalobos and E. Fereres (Editors), Principles of Agronomy for Sustainable Agriculture. Springer.

Schellnhuber, H.J., Frieler, K. and Kabat, P., 2014. The elephant, the blind, and the intersectoral intercomparison of climate impacts. Proceedings of the National Academy of Sciences, 111(9): 3225-3227.

Seidel, S.J., Palosuo, T., Thorburn, P. and Wallach, D., 2018. Towards improved calibration of crop models – Where are we now and where should we go? European Journal of Agronomy, 94: 25-35.

Semenov, M.A. and Halford, N.G., 2009. Identifying target traits and molecular mechanisms for wheat breeding under a changing climate. Journal of Experimental Botany, 60(10): 2791-2804.

Smith, P., Nkem, J., Calvin, K., Campbell, D., Cherubini, F., Grassi, G., Korotkov, V., Hoang, A.L., Lwasa, S., McElwee, P. and 4 others, 2019. Interlinkages Between Desertification, Land Degradation, Food Security and Grennhouse Gas Fluxes: Synergies, Trade-offs and Integrated Response Options. In: P.R. Shukla et al. (Editors), Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems, In press.

Soussana, J.-F., Graux, A.-I. and Tubiello, F.N., 2010. Improving the use of modelling for projections of climate change impacts on crops and pastures. Journal of Experimental Botany, 61(8): 2217-2228.

Statistisches Bundesamt, 2018. Destatis, https://www.destatis.de/.

Sutherst, R.W., Constable, F., Finlay, K., Harrington, R., Luck, J. and Zalucki, M.P., 2011. Adapting to crop pest and pathogen risks under a changing climate. WIREs Climate Change, 2: 220-237.

Tao, F., Rötter, R.P., Palosuo, T., Díaz-Ambrona, C.G.H., Mínguez, M.I., Semenov, M.A., Kersebaum, K.C., Nendel, C., Cammarano, D., Hoffmann, H. and 11 others, 2017. Designing future barley ideotypes using a crop model ensemble. European Journal of Agronomy, 82(Part A): 144-162.

Tao, F., Rötter, R.P., Palosuo, T., Gregorio Hernández Díaz-Ambrona, C., Mínguez, M.I., Semenov, M.A., Kersebaum, K.C., Nendel, C., Specka, X., Hoffmann, H. and 11 others, 2018. Contribution of crop model structure, parameters and climate projections to uncertainty in climate change impact assessments. Global Change Biology, 24(3): 1291-1307.

Tao, F., Yokozawa, M. and Zhang, Z., 2009. Modelling the impacts of weather and climate variability on crop productivity over a large area: a new process-based model development, optimization, and uncertainties analysis. Agricultural and Forest Meteorology, 149(5): 831-850.

Taylor, K.E., Stouffer, R.J. and Meehl, G.A., 2012. An Overview of CMIP5 and the Experiment Design. Bulletin of the American Meteorological Society, 93(4): 485-498.

Tebaldi, C. and Knutti, R., 2007. The use of the multi-model ensemble in probabilistic climate projections. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 365(1857): 2053-2075.

Trnka, M., Dubrovský, M. and Žalud, Z., 2004. Climate Change Impacts and Adaptation Strategies in Spring Barley Production in the Czech Republic. Climatic Change, 64(1): 227-255.

Trnka, M., Olesen, J.E., Kersebaum, K.C., Skjelvåg, A.O., Eitzinger, J., Seguin, B., Peltonen-Sainio, P., Rötter, R., Iglesias, A., Orlandini, S. and 17 others, 2011. Agroclimatic conditions in Europe under climate change. Global Change Biology, 17(7): 2298-2318.

Tuomenvirta, H., Haavisto, R., Hildén, M., Lanki, T., Luhtala, S., Meriläinen, P., Mäkinen, K., Parjanne, A., Peltonen-Sainio, P., Pilli-Sihvola, K. and 3 others, 2018. Weather and Climate Risks in Finland - National Assessment.

Turral, H., Burke, J. and Faurès, J.-M., 2011. Climate change, water and food security, Rome.

van Ittersum, M. and Rabbinge, R., 1997. Concepts in production ecology for analysis and quantification of agricultural input-output combinations. Field Crops Research, 52(3): 197-208.

van Ittersum, M.K., Cassman, K.G., Grassini, P., Wolf, J., Tittonell, P. and Hochman, Z., 2013. Yield gap analysis with local to global relevance—A review. Field Crops Research, 143(Supplement C): 4-17.

van Minnen, J.G., Alcamo, J. and Haupt, W., 2000. Deriving and Applying Response Surface Diagrams for Evaluating Climate Change Impacts on Crop Production. Climatic Change, 46(3): 317-338.

van Oijen, M. and Ewert, F., 1999. The effects of climatic variation in Europe on the yield response of spring wheat cv. Minaret to elevated CO_2 and O_3: an analysis of open-top chamber experiments by means of two crop growth simulation models. European Journal of Agronomy, 10(3): 249-264.

van Vuuren, D.P., Edmonds, J.A., Kainuma, M., Riahi, K., Thomson, A.M., Hibbard, K.A., Hurtt, G.C., Kram, T., Krey, V., Lamarque, J.-F. and 5 others, 2011. The representative concentration pathways: an overview. Climatic Change, 109: 5-31.

Voss-Fels, K.P., Stahl, A. and Hickey, L.T., 2019. Q&A: modern crop breeding for future food security. BMC Biology, 17(1): 18.

Wallach, D., Makowski, D., Jones, J.W. and Brun, F., 2014. Working with Dynamic Crop Models: Methods, Tools and Examples for Agriculture and Environment, 2nd Edition. Academic Press, London.

Wallach, D., Martre, P., Liu, B., Asseng, S., Ewert, F., Thorburn, P.J., van Ittersum, M., Aggarwal, P.K., Ahmed, M., Basso, B. and 38 others, 2018. Multimodel ensembles improve predictions of crop–environment–management interactions. Global Change Biology, 24(11): 5072-5083.

Wallach, D., Mearns, L.O., Ruane, A.C., Rötter, R.P. and Asseng, S., 2016. Lessons from climate modeling on the design and use of ensembles for crop modeling. Climatic Change, 139(3): 551-564.

Webber, H., Ewert, F., Olesen, J.E., Müller, C., Fronzek, S., Ruane, A.C., Bourgault, M., Martre, P., Ababaei, B., Bindi, M. and 21 others, 2018. Diverging importance of drought stress for maize and winter wheat in Europe. Nature Communications, 9(1): 4249.

Weigel, H.-J. and Manderscheid, R., 2012. Crop growth responses to free air CO_2 enrichment and nitrogen fertilization: rotating barley, ryegrass, sugar beet and wheat. European journal of agronomy, 43: 97-107.

Weiß, M., 2011. Future water availability in selected European catchments: a probabilistic assessment of seasonal flows under the IPCC A1B emission scenario using response surfaces. Nat. Hazards Earth Syst. Sci., 11(8): 2163-2171.

Weiß, M. and Alcamo, J., 2011. A systematic approach to assessing the sensitivity and vulnerability of water availability to climate change in Europe. Water Resour. Res., 47(2).

Wetterhall, F., Graham, L.P., Andreasson, J., Rosberg, J. and Yang, W., 2011. Using ensemble climate projections to assess probabilistic hydrological change in the Nordic region. Natural Hazards and Earth System Sciences, 11(8): 2295-2306.

Wheeler, T. and von Braun, J., 2013. Climate Change Impacts on Global Food Security. Science, 341(6145): 508-513.

White, J.W., Hoogenboom, G., Kimball, B.A. and Wall, G.W., 2011. Methodologies for simulating impacts of climate change on crop production. Field Crops Research, 124: 357-368.

White, J.W., Hunt, L.A., Boote, K.J., Jones, J.W., Koo, J., Kim, S., Porter, C.H., Wilkens, P.W. and Hoogenboom, G., 2013. Integrated description of agricultural field experiments and production: The ICASA Version 2.0 data standards. Computers and Electronics in Agriculture, 96: 1-12.

Whittaker, G., Confesor, R., Jr, Di Luzio, M. and G. Arnold, J., 2010. Detection of Overparameterization and Overfitting in an Automatic Calibration of SWAT. Transactions of the ASABE, 53(5): 1487-1499.

Willows, R. and Connell, R., 2003. Climate adaptation: Risk, uncertainty and decision making (UKCIP Tech Rep, UK Climate Impacts Programme, Oxford). 154.

YM, 2015. Climate Change Act 609/2015. Ministry of the Environment., Helsinki, Finland.

APPENDIX 1. CRITERIA FOR CLASSIFYING MODEL COMPLEXITY (TABLE 2)

The criteria for the classification as well as the classification itself were defined based on the individual model details reported in Supplement 1 of paper **I** (Table S1) and on a literature review of model references. Phenological development was not included in the analysis as all models in the ensemble simulate phenology similarly, as a function of temperature, though some may also include modifying factors such as daylength, vernalisation and other effects relating to water and nutrients.

With respect to light interception, a model is classified as having a detailed approach if it involves a representation of canopy layers, whereas a simple approach relies on the leaf area index (LAI) alone. For light utilization, simulation of photosynthesis and respiration was classified as detailed, while other approaches, mostly based on radiation use efficiency, were classified as being simple. With respect to yield formation, models simulating partitioning were classified as having a detailed approach, while approaches applying a harvest index describing the ratio of yield to total biomass were defined as being simple. The classification of the soil processes in a model was based on a rather subjective understanding of the overall detail in describing them in the model references. The central criterion was the number of soil layers in the soil profile. The definition of one or two layers was taken as an indication of the model having a simple approach while a multi-layered soil profile indicated a more detailed approach. The approach for simulating water dynamics (capacity approaches vs. Richards equation) and the emphasis put on describing nutrient dynamics were considered as supporting information for the classification. Finally, it should be noted that the terminology for describing process descriptions as being detailed or simple is used here solely as a device for obtaining an approximate sense of the relative complexity of model descriptions. It should not be interpreted as a judgement on the relative performance of the models. For detailed descriptions of the approaches applied for simulating the different processes in each model see Supplement 1 in paper **I.**

ERRATUM

Paper I, Supplementary material
http://www.int-res.com/articles/suppl/c065p087_supp.pdf **:**

Figure S3 shows the results of spring wheat, already shown in Figure 7 of the main article. The correct figure for winter wheat is shown below.

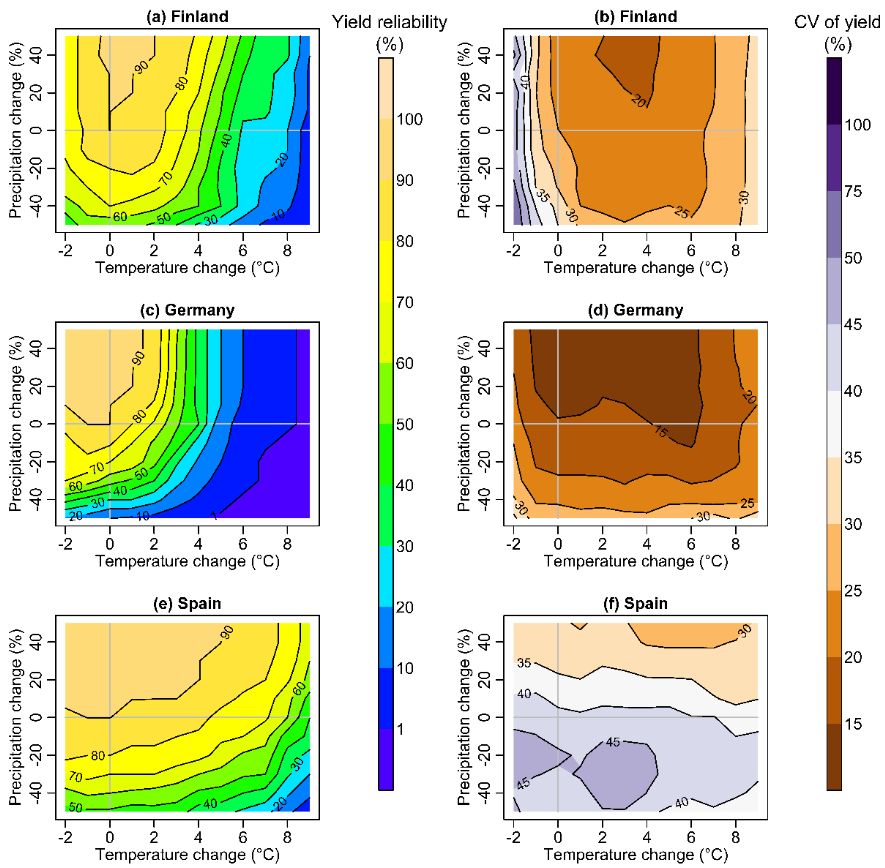

Figure S3. Ensemble medians of yield reliability, defined as the percentage of years when DM grain yield (kg ha^{-1}) is above the 10th percentile of the baseline yield (left-hand panels), and of coefficients of variation (CV) of annual yields (right-hand panels), for winter wheat under changes in temperature and precipitation relative to the 1981–2010 baseline for 26 crop models at Jokioinen, Finland (a and b), Dikopshof, Germany (c and d) and Lleida, Spain (e and f).